

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 19, 2002, 15:23:43 ; Search time 26.01 Seconds
(without alignments)
4330.246 Million cell updates/sec

Title: US-08-212-185-12

Sequence: 1 MAQWQLOQLDRYLKQLHQ.....QFSLTFMDLTSECATSPM 770

Scoring table: Gapop 60.0, Gapext 60.0

Searched: 473505 seqs, 146272329 residues

Word size: 0

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

1: SPREMBL_17:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mhc:*
8: sp_mammal:*
9: sp_organelle:*
10: sp_phase:*
11: sp_plant:*
12: sp_proteus:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	532	69.1	770	4	014916 homo sapien
2	473	61.4	769	4	09BM54 homo sapien
3	271	35.2	769	13	09PVX8 xenopus lae
4	145	18.8	163	6	09N145 macaca mula
5	63	8.2	394	13	09DDJ8 brachydanio
6	63	8.2	806	13	093599 brachydanio
7	59	7.7	767	13	013133 oncorhynch
8	56	7.3	56	6	09N034 bos taurus
9	51	6.6	51	11	099ML3 mus musculu
10	43	5.6	43	11	09QVR4 rattus sp.
11	38	4.9	51	4	09BMX2 oncorhynch
12	19	2.5	754	13	013132 oncorhynch
13	12	1.6	43	11	09QVR3 rattus sp.
14	12	1.6	156	4	09UDL5 mus musculu
15	12	1.6	712	11	099K94 mus musculu
16	12	1.6	749	11	09QXK0 rattus norv
17	12	1.6	749	11	09D323 mus musculu
18	12	1.6	754	13	013131 oncorhynch
19	11	1.4	155	4	09UDL4 homo sapien

20	11	1.4	195	6	09GKY9 sus scrofa
21	11	1.4	749	13	093598 brachydanio
22	9	1.2	44	5	09XXE3 caenorhabdi
23	9	1.2	72	5	018168 caenorhabdi
24	9	1.2	732	5	09VY71 drosophila
25	9	1.2	922	11	09QXJ2 mus musculu
26	9	1.2	925	11	09QCE4 mus musculu
27	8	1.0	85	6	029340 sus scrofa
28	8	1.0	111	11	070406 rattus norv
29	8	1.0	141	11	070429 rattus norv
30	8	1.0	177	6	09N0E5 bos taurus
31	8	1.0	293	2	09KRP8 vibrio chol
32	8	1.0	348	11	099K06 mus musculu
33	8	1.0	515	10	09FKI8 arabidopsis
34	8	1.0	581	2	087594 streptomyce
35	8	1.0	588	2	09RJ58 streptomyce
36	8	1.0	592	13	09PMP7 gallus gall
37	8	1.0	679	4	09BQD2 anopheles g
38	8	1.0	722	5	097164 halobacteri
39	8	1.0	728	1	09HOB3 oncorhynch
40	8	1.0	784	13	09D374 mus musculu
41	8	1.0	786	11	09JXM1 gallus gall
42	8	1.0	787	13	093378 kapoti's sa
43	8	1.0	843	12	088944 avian rotav
44	8	1.0	1088	12	055590 porcine rot
45	8	1.0	1088	12	085036 porcine rot

ALIGNMENTS

RESULT 1
ID 014916 PRELIMINARY: PRT: 770 AA.
AC 014916:
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-AUG-1999 (TREMBLrel. 11, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE TRANSCRIPTION FACTOR (SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION
3).
GN STAT3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98296260; PubMed=9630560;
RA Della Pietra L., Bressan A., Pezzotti A., Serlupi-Crescenzi O.;
RT "Highly conserved amino-acid sequence between murine STAT3 and a
revised human STAT3 sequence.";
RL Gene 213:119-124(1998).
RN [2]
RP SEQUENCE OF 564-704 FROM N.A.
RC TISSUE=HEPATOMA;
RA Della Pietra L., Bressan A., Pezzotti A.R., Serlupi-Crescenzi O.;
RL Submitted (Oct-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: AJ012463; CAI10032.1;
DR EMBL: AF029311; AAB84254.1;
DR HSSP: P42227; BGL.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PSS0001; SH2; 2.
SQ SEQUENCE 770 AA: 88067 MM: 6C00632211C8012D CRC64:

Query Match 69.1%; Score 532; DB 4; Length 770;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	26	FMPELROFLAPMLIESOMAAVAAKSESHATVFNILGELIDDOOYSRFLQSSNVLYONHLR	85
Db	26	FMPELROFLAPMLIESODMAVAASESHATVFNILNGEIDDOOYSRFLQSSNVLYONHLR	85
QY	86	IKQFLQSHRYLEKPEMEIARIIVARCLMEFSRLLOTAATAAODGOGAHPHTAAVTEKOOMLE	145
Db	86	IKQFLQSHRYLEKPEMEIARIIVARCLMEFSRLLOTAATAAODGOGAHPHTAAVTEKOOMLE	145
QY	146	OHLODYKRRVODLEQKKKVVENLDDDFEYKTLKSQGMQDLNGNNOYSTRQKKOOLEQ	205
Db	146	QHLODYKRRVODLEQKKKVVENLDDDFEYKTLKSQGMQDLNGNNOYSTRQKKOOLQO	205
QY	206	MUTALDQRRSIVSELGGLLSAMEYOKTLTDEELADMKRPEIACIGCPNLCIDRLN	265
Db	206	MUTALDQRRRSIVSELGGLLSAMEYOKTLTDEELADMKRROQIACIGCPNLCIDRLN	265
QY	266	WITSLSAESLOLTROOIKKLEFLOOKVSYKQDPYVQHRPMLEERIVELFRILMKSAAVVER	325
Db	266	WITSLSAESLOLTROOIKKLEFLOOKVSYKQDPYVQHRPMLEERIVELFRILMKSAAVVER	325
QY	326	OPCMHMHDRPLVITKTOGQFTTKRLLVKKPELNYOLKIVCIDKDSGDVAALRGSRKEN	385
Db	326	OPCMHMHDRPLVITKTOGFTTKRLLVKKPELNYOLKIVCIDKDSGDVAALRGSRKEN	385
QY	386	ILGTNTKVMNNEESNNGSLAEFKHLTLREORCGNGRANCASLIYTELHLITFETEV	445
Db	386	ILGTNTKVMNNEESNNGSLAEFKHLTLREORCGNGRANCASLIYTELHLITFETEV	445
QY	446	YHOGJKTIDLEHSHLPVVYISNICOMPANAMASILYNNLTNNPKNVNFETKPICTMDOVA	505
Db	446	YHOGJKTIDLEHSHLPVVYISNICOMPANAMASILYNNLTNNPKNVNFETKPICTMDOVA	505
QY	506	EVLNMOFSSSTTKRGSLIEQLTTLAEKLLGPGVNSGCOITWAKFCENMAGKPSFWWL	565
Db	506	EVLNMOFSSSTTKRGSLIEQLTTLAEKLLGPGVNSGCOITWAKFCENMAGKPSFWWL	565
QY	566	DNIIDLVKKYLALMNEGYINGFISKEPERAILSTKRPGRFLRPFSSSEKGGVTFTWVE	625
Db	566	DNIIDLVKKYLALMNEGYINGFISKEPERAILSTKRPGRFTLLRPFSSSEKGGVTFTWVE	625
QY	626	KDISKTOIQSVEPYTKOOLNNMSFAETIIMGYKIMDATNIIIVSFLVLYPDIFKEAFCG	685
Db	626	KDISKTOIQSVEPYTKOOLNNMSFAETIIMGYKIMDATNIIIVSFLVLYPDIFKEAFCG	685
QY	686	YCRPESQEPHADPCSAPYIKTKFICVTPPTGCSNTIIDLPMSPPTDLSLMOFGNNGGAE	745
Db	686	YCRPESQEPHADPCSAPYIKTKFICVTPPTGCSNTIIDLPMSPPTDLSLMOFGNNGGAE	745
QY	746	PSAGGOFELTFDDM 759	
Db	746	PSAGGOFESTLFDPM 759	
RESULT 2			
Q9BM54		PRELIMINARY;	PRF: 769 AA.
AC	Q9BM54;		
DT	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 3.		
OS	Homo sapiens (Human)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
NCBI	NCBI_Taxid=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=KIDNEY ADENOCARCINOMA;		
RL	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.		
SR	EMBL: BC000627; AAH00627.1; -.		
SR	SEQUENCE 769 AA: 87980 MW: A374A32AB9D28077 CRC64:		

Query Match	61.4%	Score 473	DB 4	Length 769	
Best Local Similarity	99.7%	Pred. No. 0			
Matches 673	Conservative	0	Mismatches 2	Indels 0	Gaps 0
QY 26	FPMLRQFLAMISODPAYAASKESHATLVFHNILGLEDIDQYSRFLDESNTLYOHNLRR	85			
Db 26	FPMLRQFLAPMIESQDMAYAAKSESHATLVFHNILGLEDIDQYSRFLDESNTLYOHNLRR	85			
QY 86	IKQFLQSHYLEKPEMEIARIIVAKCIWEEESHLLQTAATAAQOGQAMHPRAAVTEKQUMLE	145			
Db 86	IKQFLQSHYLEKPEMEIARIIVAKCIWEEESHLLQTAATAAQOGQAMHPRAAVTEKQUMLE	145			
QY 146	OHLDQVRRKVDLEQKMKVVENLQDDPFENKTKLSQSGMODLNKNNQSVTRQKKQOLEQ	205			
Db 146	OHLDQVRRKVDLEQKMKVVENLQDDPFENKTKLSQSGMODLNKNNQSVTRQKKQOLEQ	205			
QY 206	MLTALDQMRISIVSELACLLSMEYVQKTLTDEELADMKRPEIACIGSPNICIDRLLEN	265			
Db 206	MLTALDQMRISIVSELACLLSMEYVQKTLTDEELADMKRROUICIGSPNICIDRLLEN	265			
QY 266	WITSLSAQLOTRQOIKKLELEQGVSKGDPYVGRHPLMERIYELFENLKKSAFYVER	325			
Db 266	WITSLSAQLOTRQOIKKLELEQGVSKGDPYVGRHPLMERIYELFENLKKSAFYVER	325			
QY 326	QPCMPMHDPRLVITKGVQFTTKVRLVYKFPRELNYQLIKVCIDKDSGDVAALRGSKKFN	385			
Db 326	QPCMPMHDPRLVITKGVQFTTKVRLVYKFPRELNYQLIKVCIDKDSGDVAALRGSKKFN	385			
QY 386	ILGTNTKVMNNEESNGSLSAFFKRLTLREDCGNGCANCDASLIYVEELHLTFETFEV	445			
Db 386	ILGTNTKVMNNEESNGSLSAFFKRLTLREDCGNGCANCDASLIYVEELHLTFETFEV	445			
QY 446	YHOGKLIQLETHSLPVLVYISNICQMPNMAWILWYNMLTNNDKNVNFPTKPIGTMDQVA	505			
Db 446	YHOGKLIQLETHSLPVLVYISNICQMPNMAWILWYNMLTNNDKNVNFPTKPIGTMDQVA	505			
QY 506	EVLWSQSFSTTKRGSLTQLTTLAKKLLGPGVNYGGCOITMAKPFCKENNAGGFSFWML	565			
Db 506	EVLWSQSFSTTKRGSLTQLTTLAKKLLGPGVNYGGCOITMAKPFCKENNAGGFSFWML	565			
QY 566	DNIIIDLVRKYILALNNEGIMGFIKEREERALLSTRPGCTFLRLRESSEKSGGVFTWVE	625			
Db 566	DNIIIDLVRKYILALNNEGIMGFIKEREERALLSTRPGCTFLRLRESSEKSGGVFTWVE	625			
QY 626	KDISKTOIQSEPTTKQOLNNNSFAETIMGKINDATNLLVSPLYLYXPDIPKEARECK	685			
Db 626	KDISKTOIQSEPTTKQOLNNNSFAETIMGKINDATNLLVSPLYLYXPDIPKEARECK	685			
QY 686	YCRPESQEHPEADPG 700				
Db 686	YCRPESQEHPEADPG 700				
RESULT 3					
Q9PVX8	PRELIMINARY:	PRT:	769	AA.	
AC Q9PVX8:					
DT 01-MAY-2000 (TREMBLrel. 13, Created)					
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)					
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)					
DE START 3.					
GN START 3.					
OS Xenopus laevis (African clawed frog).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;					
OC Xenopodinae; Xenopus.					
OX NCBI_TaxID=8355;					
RM [1]					
RP SEQUENCE FROM N.A.					
RA Nishinakamura R., Matsunoto Y., Matsuda T., Arizumi T., Heike T.,					
RA Asashima M., Yokota T.;					
RT *Activation of Stat3 by cytokine receptor gp130 ventralizes Xenopus					

Oy 249 IACIGPPNICDLRLNNWITSLAESQLQTRQIKLLEELQOKVSYKDPYVHRPMLER 300
|||||
|||
Db 19 IACIGPPNICDLRLNNWITSLAESQLQTRQIKLLEELQOKVSYKDPYVHRPMLER 78

RA Oates A.C.?
RL Thesis (1998), University of Melbourne, Australia.
DR EMBL; AJ005693; CAA06677.1; -.

DR HSSP: P42227; 1BG1.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF00117; STAT; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS0001; SH2; 1.
 SO SEQUENCE 806 AA; 92151 MW; 7ABCAEAA01C3C942 CRC64;

Query Match 8.2%; Score 63; DB 13; Length 806;
 Best Local Similarity 100.0%; Pred. No. 1.5e-57;
 Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 537 VNYSGQITWAKFKENNAKGFSFWVLNDITLVKRYIALALNMGYIMGFISKERERA 596
 DB 538 VNYSGQITWAKFKENNAKGFSFWVLNDITLVKRYIALALNMGYIMGFISKERERA 597
 OY 597 ILS 599
 DB 598 ILS 600

RESULT 7
 O13133 PRELIMINARY: PRT: 767 AA.
 AC 013133:
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE STAT3.
 GN RBTSTAT3.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 CC Prochordopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Johnson M.C., Mourich D.V., Leong J.C.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U60333; AAB60926.1; -.
 DR HSSP: P42227; 1BG1.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF01017; STAT; 1.
 SO SEQUENCE 767 AA; 87816 MW; FED97740C74C3798 CRC64;

Query Match 7.7%; Score 59; DB 13; Length 767;
 Best Local Similarity 100.0%; Pred. No. 2.5e-53;
 Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 427 DASLVTBELHLITRETVYHOGKIDLETHSLPVVVISNICOMPANASILMYMLTN 485
 DB 428 DASLVTBELHLITRETVYHOGKIDLETHSLPVVVISNICOMPANASILMYMLTN 486

RESULT 8
 O9N0J4 PRELIMINARY: PRT: 56 AA.
 AC 09N0J4:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 3, STAT3 (FRAGMENT).
 GN STAT3.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Molenar A., Wheeler T.T., McCracken J.Y., Seyfert H.M.;
 RT "The STAT3-encoding gene resides within the 40 kbp gap between the
 RT STAT5A- and STAT5B-encoding genes in cattle."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ276489; CAB93140.1; -.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF01017; STAT; 1.
 FT NON_TER 1
 FT NON_TER 56
 SO SEQUENCE 56 AA; 6057 MW; 4450AC5D0DB6D4A CRC64;

Query Match 7.3%; Score 56; DB 6; Length 56;
 Best Local Similarity 100.0%; Pred. No. 3.6e-51;
 Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 372 SGGVALALGSKRKNILGINTVMMNNEESNNGSLSAEFKHLTLREORCGNGGRAND 427
 DB 1 SGGVALALGSKRKNILGINTVMMNNEESNNGSLSAEFKHLTLREORCGNGGRAND 56

RESULT 9
 O99ML3 PRELIMINARY: PRT: 51 AA.
 AC 099ML3:
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE STAT3 (FRAGMENT).
 GN STAT3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shao H., Tweardy D.J.;
 RT "Murine Stat3 gene exon 22 to exon 23."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF332507; AAK17195.1; -.
 DR NON_TER 1
 FT NON_TER 51
 SO SEQUENCE 51 AA; 5379 MW; CE2B44AD9932D12 CRC64;

Query Match 6.6%; Score 51; DB 11; Length 51;
 Best Local Similarity 100.0%; Pred. No. 6.9e-46;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 702 AAPYLKTFKFCVPTTGCSNTIDLPMSPTLDLSLQFGNNGEGAEPSAGGF 752
 DB 1 AAPYLKTFKFCVPTTGCSNTIDLPMSPTLDLSLQFGNNGEGAEPSAGGF 51

RESULT 10
 O9QVR4 PRELIMINARY: PRT: 43 AA.
 AC 09QVR4:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE INTERLEUKIN-6 RESPONSE ELEMENT BINDING PROTEIN (FRAGMENT).
 OS Rattus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10118;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLIN=95398099; Pubmed=7545365;
 RA Ripberger J., Fritz S., Richter K., Dreier B., Schneider K.,
 Lochner K., Marschalek R., Hocke G., Lottspeich F., Fey G.H.;

```

RT      *isolation of two interleukin-6 response element binding proteins from
RT      acute phase rat livers.*
RL      Ann. N. Y. Acad. Sci. 762:252-260(1995).
DR      HSSP: P42227: 1BG1.
DR      InterPro: IPR000980: SH2.
DR      Pfam: PF00017: SH2: 1.
DR      PROSITE: PS50001: SH2: 1.
DR      PROSITE: PS50001: SH2: 1.
SQ      SEQUENCE 43 AA: 4868 MW: 0C2F14721B863798 CRC64:

Query Match
Best Local Similarity 100.0%; Score 43; DB 11; Length 43;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

OY      581 NEGIMGFISKERERAILSTKPGCTFLLRFSESSKEGCVFTW 623
          |||||||
DB      1 NEGIMGFISKERERAILSTKPGCTFLLRFSESSKEGCVFTW 43

RESULT 11
ID      Q9BXH2      PRELIMINARY:      PRT:      51 AA.
AC      Q9BXH2;
DT      01-JUN-2001 (TREMBLrel. 17, Created)
DT      01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE      INTERLEUKIN-6, RESPONSE ELEMENT BINDING PROTEIN (FRAGMENT).
DE      STAT3 (FRAGMENT).
GN      STAT3.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Shao H., Tweardy D.J.:
RT      "Human STAT3 gene exon 22 to exon 23."
RL      Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AF332508; AAK17196.1; -.
FT      NON_TER      1      51
FT      NON_TER      1      51
SQ      SEQUENCE 51 AA: 5375 MW: CE2B444D8AFC3D12 CRC64:

Query Match
Best Local Similarity 100.0%; Score 38; DB 4; Length 51;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

OY      702 AAPYLTAKTFCVPTTCSNTIDLPMSPTLDLSLMQFCN 739
          |||||||
DB      1 AAPYLTAKTFCVPTTCSNTIDLPMSPTLDLSLMQFCN 38

RESULT 12
ID      O13132      PRELIMINARY:      PRT:      754 AA.
AC      O13132;
DT      01-JUL-1997 (TREMBLrel. 04, Created)
DT      01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DE      STAT1-2.
DE      STAT1-2.
GN      RARSSTAT1-2.
OS      Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC      Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX      NCBI_TaxID=8022;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Johnson M.C., Mourich D.V., Leong J.C.:
RL      Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR      EMBL: U60332; AAB60925.1; -.
DR      HSSP: P42224: 1BF5.
RT      InterPro: IPR000980; SH2.

```

```

DR      InterPro: IPR001217; STAT.
DR      Pfam: PF00017: SH2: 1.
DR      Pfam: PF01017: STAT: 1.
DR      SMART: SM00252: SH2: 1.
DR      PROSITE: PS50001: SH2: 1.
DR      PROSITE: PS50001: SH2: 1.
SQ      SEQUENCE 754 AA: 87147 MW: 340B9645EA040142 CRC64:

Query Match
Best Local Similarity 100.0%; Score 19; DB 13; Length 754;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

OY      382 RKENILCTNTKVMNMEESN 400
          |||||||
DB      373 RKENILCTNTKVMNMEESN 391

RESULT 13
ID      Q9QVR3      PRELIMINARY:      PRT:      43 AA.
AC      Q9QVR3;
DT      01-MAY-2000 (TREMBLrel. 13, Created)
DT      01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE      INTERLEUKIN-6, RESPONSE ELEMENT BINDING PROTEIN (FRAGMENT).
DE      INTERLEUKIN-6, RESPONSE ELEMENT BINDING PROTEIN (FRAGMENT).
OS      Rattus sp.
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10118;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Ripberger J., Fritz S., Richter K., Dreier B., Schneider K.,
RT      Lochner K., Marschalek R., Hocke G., Lotzpeich F., Fey G.H.:
RT      "Isolation of two interleukin-6 response element binding proteins from
RT      acute phase rat livers."
RL      Ann. N. Y. Acad. Sci. 762:252-260(1995).
DR      HSSP: P42224: 1BF5.
DR      InterPro: IPR000980; SH2.
DR      Pfam: PF00017: SH2: 1.
DR      PROSITE: PS50001: SH2: 1.
SQ      SEQUENCE 43 AA: 4937 MW: 5D710A6749125C3B CRC64:

Query Match
Best Local Similarity 100.0%; Score 12; DB 11; Length 43;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

OY      585 IMGFISKEKERA 596
          |||||||
DB      5 IMGFISKEKERA 16

RESULT 14
ID      Q9UDL5      PRELIMINARY:      PRT:      156 AA.
AC      Q9UDL5;
DT      01-MAY-2000 (TREMBLrel. 13, Created)
DT      01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE      INTERFERON-STIMULATED GENE FACTOR 3 ALPHA 91/84 KDA PROTEIN
DE      (FRAGMENTS).
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      MEDLINE=92346719; PubMed=1638633;
RX      Fu X.Y.:
RT      "A transcription factor with SH2 and SH3 domains is directly activated
RT      by an interferon alpha-induced cytoplasmic protein tyrosine
RT      kinase(s).".

```

RL Cell 70:323-335(1992).
 DR HSSP: P42224; 1BF5.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF01017; STAT; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS50001; SH2; 1.
 FT NON_TER 1
 FT NON_CONS 54
 FT NON_TER 156
 SQ SEQUENCE 156 AA; 18004 MW; 4D6E4D7D353C0AAA CRC64;

Query Match 1.6%; Score 12; DB 4; Length 156;
 Best Local Similarity 100.0%; Pred. No. 0.00051;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 585 IMGFISKERERA 596
 |||||
 DB 60 IMGFISKERERA 71

RESULT 15
 099K94 PRELIMINARY: PRT: 712 AA.
 AC 099K94;
 DT 01-JUN-2001 (TREMBLER, 17, Created)
 DT 01-JUN-2001 (TREMBLER, 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLER, 17, Last annotation update)
 DE SIMILAR TO SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC004808; AAH04808.1; -;
 SQ SEQUENCE 712 AA; 83106 MW; D62081709638CDA5 CRC64;

Query Match 1.6%; Score 12; DB 11; Length 712;
 Best Local Similarity 100.0%; Pred. No. 0.002;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 585 IMGFISKERERA 596
 |||||
 DB 578 IMGFISKERERA 589

Search completed: March 19, 2002, 15:27:46
 Job time: 243 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 19, 2002, 15:22:48 ; Search time 14.48 Seconds

(Without alignments)
1949.718 Million cell updates/sec

Title: US-08-212-185-12

Sequence: 1 MAQWNLQQLDRLRYLKQLHQ.....QFSLTLPMDLTSECATSPM 770

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 100059 seqs, 36664827 residues

Word size : 0

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	543	70.5	770	STA3_MOUSE	P42227 mus musculu
2	459	59.6	770	STA3_RAT	P52631 rattus norv
3	221	28.7	770	STA3_HUMAN	P40763 homo-sapien
4	12	1.6	749	STA1_MOUSE	P42225 mus musculu
5	12	1.6	750	STA1_HUMAN	P42224 homo sapien
6	11	1.4	748	STA4_HUMAN	Q14765 homo sapien
7	11	1.4	749	STA4_MOUSE	P42228 mus musculu
8	11	1.4	851	STA2_HUMAN	P52630 homo sapien
9	9	1.2	864	STA2_MOUSE	Q02799 sus scrofa
10	9	1.2	923	STA2_MOUSE	Q02799 sus musculu
11	8	1.0	458	HDNO_ARTOX	P08159 arthobacte
12	8	1.0	663	MNEL_YEAST	P24720 saccharomyc
13	8	1.0	761	STA1_DROME	Q24151 drosophila
14	8	1.0	766	STA5_HUMAN	P51692 homo sapien
15	8	1.0	766	STA5_MOUSE	P42232 mus musculu
16	8	1.0	766	STA5_RAT	P52632 rattus norv
17	8	1.0	787	STA5_BOVIN	Q02633 bos taurus
18	8	1.0	787	STA5_MOUSE	Q02633 sus scrofa
19	8	1.0	793	STA5_MOUSE	P42230 mus musculu
20	8	1.0	793	STA5_RAT	Q62771 rattus norv
21	8	1.0	794	STA5_BOVIN	Q05115 bos taurus
22	8	1.0	794	STA5_HUMAN	P42229 homo sapien
23	8	1.0	794	STA5_SHEEP	P42231 ovis aries
24	8	1.0	799	STA5_MOUSE	Q02633 sus scrofa
25	8	1.0	837	STA6_MOUSE	Q02633 mus musculu
26	8	1.0	847	STA6_HUMAN	P42226 homo sapien
27	8	1.0	1088	RRPO_ROTBR	P17468 bovine rota
28	8	1.0	1088	RRPO_ROTBU	P17468 bovine rota
29	8	1.0	1088	RRPO_ROTGS	P17469 porcine rota
30	8	1.0	1088	RRPO_ROTSL	P22678 simian 11 r
31	7	0.9	74	YD77_MCPN	Q06053 mycoplasma
32	7	0.9	75	RL29_MYCBO	Q06053 mycobacteri
33	7	0.9	77	RL29_MYCTU	P95057 mycobacteri

ALIGNMENTS

RESULT	ID	STANDARD	PRT	770 AA	
1	STA3_MOUSE				
AC	STA3_MOUSE				
AD	P42227				
DT	01-NOV-1995 (Rel. 32, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	20-AUG-2001 (Rel. 40, Last annotation update)				
DE	SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 3 (ACUTE-PHASE RESPONSE FACTOR).				
GN	STAT3 OR APRF.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID:10090;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 154-158; 181-185 AND 632-640.				
RC	STRAIN-BALB/C; TISSUE=Liver;				
RX	MEDLINE=94208062; PubMed=7512451;				
RA	Akira S., Nishio Y., Inoue M., Wang X.J., Shi W., Matsusaka T.,				
RA	Yoshida K., Sudo T., Naruto M., Kishimoto T.,				
RT	"Molecular cloning of APRF, a novel IFN-gamma-stimulated gene factor 3 p91-				
RT	related transcription factor involved in the gp130-mediated signaling				
RT	pathway.";				
RL	Cell 77:63-71(1994).				
RP	[2]				
RC	SEQUENCE FROM N.A.				
RX	TISSUE=Thymus;				
RA	MEDLINE=94188718; PubMed=8140422;				
RA	Zhong Z., Wen Z., Darnell J.E., Jr.;				
RT	"Stat3: a STAT family member activated by tyrosine phosphorylation in				
RT	response to epidermal growth factor and interleukin-6.";				
RL	Science 264:95-98(1994).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Brain;				
RX	MEDLINE=95014185; PubMed=7523373;				
RA	Raz R., Durbin J.E., Levy D.E.;				
RT	"Acute phase response factor and additional members of the				
RT	interferon-stimulated gene factor 3 family integrate diverse signals				
RT	from cytokines, interferons, and growth factors.";				
RL	J. Biol. Chem. 269:24391-24395(1994).				
RN	[4]				
RP	SEQUENCE FROM N.A. (STAT3B).				
RC	STRAIN-BALB/C; AND B6; TISSUE=Liver;				
RX	MEDLINE=96016116; PubMed=7568080;				
RA	Schaefer T.S., Sanders L.K., Nathans D.;				
RT	"Cooperative transcriptional activity of Jun and Stat3 beta, a short				
RT	form of Stat3.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 92:9097-9101(1995).				
RN	[5]				
RP	PHOSPHORYLATION OF SER-727, AND MUTAGENESIS.				
RX	MEDLINE=95354205; PubMed=7543024;				
RA	Wen Z., Zhong Z., Darnell J.E., Jr.;				
RT	"Maximal activation of transcription by Stat1 and Stat3 requires both				
RT	tyrosine and serine phosphorylation.";				
RL	Cell 82:241-250(1995).				

34	7	0.9	94	1	SPDA_STRLI	P22407 streptomyc
35	7	0.9	128	1	VF70_AQUAE	O67514 aquifex ae
36	7	0.9	136	1	SR14_HUMAN	P37108 homo sapien
37	7	0.9	161	1	CC31_YEAST	P06704 saccharomyc
38	7	0.9	170	1	Y586_METJA	O58006 methanococ
39	7	0.9	173	1	PSAL_MASJA	O31126 masliogoclad
40	7	0.9	193	1	E631_DROME	P48593 drosophila
41	7	0.9	205	1	MD21_HUMAN	Q13257 homo sapien
42	7	0.9	205	1	MD21_MOUSE	Q92105 mus musculu
43	7	0.9	209	1	YRBK_HABIN	P45075 haemophilus
44	7	0.9	221	1	ERG2_MAGGR	P33281 magnaporthe
45	7	0.9	238	1	AGRA_STAAU	P13131 staphylococ

```

RN      [6] X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 136-716.  
RX      MEDLINE=98334373; PubMed=9671298;  
RA      Becker S., Groner B., Mueller C.W.;  
RT      "Three-dimensional structure of the Statbeta homodimer bound to  
       DNA.";  
RL      Nature 394:145-151(1998).  
CC      -I- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE INTERLEUKIN-6  
       (IL-6)-RESPONSIVE ELEMENT IDENTIFIED IN THE PROMOTERS OF VARIOUS  
       ACUTE-PHASE PROTEIN GENES. STAT3B INTERACTS WITH THE N-TERMINAL  
       PART OF C-JUN TO ACTIVATE SUCH PROMOTERS IN A COOPERATIVE WAY.  
CC      -I- PATHWAY: INVOLVED IN THE GPI30-MEDIATED SIGNALING PATHWAY.  
CC      -I- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY  
       MEMBER (AT LEAST STAT1) (BY SIMILARITY).  
CC      -I- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN  
       RESPONSE TO PHOSPHORYLATION.  
CC      -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS: STAT3A (SHOWN HERE) AND STAT3B:  
       ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC      -I- TISSUE SPECIFICITY: STAT3A IS SEEN IN THE LIVER, SPLEEN, AND  
       KIDNEY. STAT3B IS ALSO DETECTED IN THE LIVER, ALTHOUGH IN A MUCH  
       LESS ABUNDANT MANNER.  
CC      -I- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IL-6, IL-11, CNF,  
       LIF, CSF-1, EGF, PDGF, IFN-ALPHA AND OSM. SERINE PHOSPHORYLTION  
       IS IMPORTANT FOR THE FORMATION OF STABLE DNA-BINDING STAT3  
       HOMODIMERS AND MAXIMAL TRANSCRIPTONAL ACTIVITY (BY SIMILARITY).  
CC      -I- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.  
CC      -I- SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
-----  
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
       between the Swiss Institute of Bioinformatics and the EMBL Outstation  
       at the European Bioinformatics Institute. There are no restrictions on its  
       use by non-profit institutions as long as its content is in no way  
       modified and this statement is not removed. Usage by and for commercial  
       entities requires a license agreement (see http://www.isb-sib.ch/announce/  
       or send an email to license@isb-sib.ch).  
-----  
DR      EMBL: L29278; AAA37254.1; -.  
DR      EMBL: U06922; AAAl9452.1; -.  
DR      EMBL: U08378; AAA56668.1; -.  
DR      EMBL: U30709; AAC52612.1; -.  
DR      PDB: 1BG1; I3-Jan-99.  
DR      TRNSFCAC: T01574; -.  
DR      MGD: MG1:I03038; stat3.  
DR      InterPro: IPRO00980; SH2.  
DR      Interpro: IPR001217; STAT.  
DR      Pfam: PF00017; SH2_1.  
DR      SMART: PF01017; STAT_1.  
DR      SMART: SM00252; SH2_1.  
DR      PROSITE: PS50001; SH2; 1.  
KW      Transcription regulation; DNA-binding; Nuclear protein; SH2 domain;  
KW      Phosphorylation; Alternative splicing; 3D-structure.  
FT      MOD_RES          580   670  
FT      MOD_DOMAIN     705  
FT      MOD_RES          705  
FT      MOD_RES          705  
FT      MOD_RES          705  
FT      MOD_RES          705  
FT      MOD_RES          705  
FT      MOD_RES          705  
FT      MOD_RES          705  
FT      VARSPPLIC        716   770  
FT      FT              716   770  
FT      MUTAGEN         701   701  
FT      VARIANT         727   727  
FT      CONFLICT        16    16  
FT      CONFLICT        25    25  
FT      CONFLICT        394    394  
SO      SEQUENCE        770 AA; 88053 MM; 6C00626711C8012D CRC64;
```

Db	26	FPMELRQFLAPWIESQDMAYAAASKESHATLVFHNHLLGE	LDQYSPHJUDSESVLYQHNLRK	85
Qy	86	IKOFLQSRYLEKPMELIARIVARCLMEESRLLOTAATAOOGG	ANHPAAVTEKQOMLE	145
Db	86	IKOFLQSRYLEKPMELIARIVARCLMEESRLLOTAATAOOGG	ANHPAAVTEKQOMLE	145
Qy	146	OHLDQVRKRVODLPQKMKVVENLDDDFPNFKTLKSGQDMODL	KGNNOSYTRQKMOOLEQ	205
Db	146	OHLDQVRKRVODLEBKMKVVENLDDDFPNFKTLKSGQDMODL	KGNNOSYTRQKMOOLEQ	205
Qy	206	MILTALDQMRRSIVSELGALLSAMEYVOKTTLTDEELADMR	KRPPELACIGCPNIIQDLREN	265
Db	206	MILTALDQMRRSIVSELGALLSAMEYVOKTTLTDEELADMR	KRPQJACIGCPNIIQDLREN	265
Qy	266	WTTSLAESOLOTRQOIKKLEELQOKVSKGSPITVOIHRPML	EEERIVELFRNLMSAPFYER	325
Db	266	WTTSLAESOLOTRQOIKKLEELQOKVSKGPIVOIHRPML	EERIVELFRNLMSAPFYER	325
Qy	326	QPCMPMHDPRLPVIAIKTVQVFTTKVNLVKRPPELWYOLK	IKVCJIKKDSQDIAALKGSKKFN	385
Db	326	QPCMPMHDPRLPVIAIKTVQVFTTKVNLVKRPPELWYOLK	IKVCIDKDSQDAVALRGSKKFN	385
Qy	386	ILGTTTKVMNNEESNGSLAEFKHLLTIREORCNGGRANC	DASLIYTBELHLITFETEV	445
Db	386	ILGTTTKVMNNEESNGSLAEFKHLLTIREORCNGGRANC	DASLIYTBELHLITFETEV	445
Qy	446	YHOGAKIDLETHSLPVAVVISNICOMPNAASILWYNNLT	NPNKRVNFFTRKPIGTMQVA	505
Db	446	YHOGAKIDLETHSLPVAVVISNICOMPNAASILWYNNLT	NPNKRVNFFTRKPIGTMQVA	505
Qy	506	EYLSWQFSSTTKRGSLIEQLTTLAEKLLGPGVNSGCOIT	MAFKCKENMAGKGFSEVWL	565
Db	506	EYLSWQFSSTTKRGSLIEQLTTLAEKLLGPGVNSGCOIT	MAFKCKENMAGKGFSEVWL	565
Qy	566	DNIIDLVKKYLIALMNEGYIMGFISKEKEERAILSTKPR	GTFLFPFSESSKEGGVTFPWE	625
Db	566	DNIIDLVKKYLIALMNEGYIMGFISKEKEERAILSTKPR	GTFLFPFSESSKEGGVTFPWE	625
Qy	626	KDISKTOIQSEVPTKQQLNNMSFAEIIIMYKIMDATN	ILVSPLYLYPDIPKEEAFGK	685
Db	626	KDISKTOIQSEVPTKQQLNNMSFAEIIIMYKIMDATN	ILVSPLYLYPDIPKEEAFGK	685
Qy	686	YCRPESQHPREAPDESAPARYLKTFCITVPTTCSNT	IDLPMSPRPLDLSLQFGNCGAE	745
Db	686	YCRPESQHPREAPDESAPARYLKTFCITVPTTCSNT	IDLPMSPRPLDLSLQFGNCGAE	745
Qy	746	PSAGQCFESLTFDMDLTSECATSPM	770	
Db	746	PSAGQCFESLTFDMDLTSECATSPM	770	
RESULT 2				
STAJ3_RAT	STAJ3_RAT	STANDARD:	PRT:	770 AA.
AC	P52631:			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 3.			
GN	STAT3.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euteleia; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	SEQUENCE FROM N.A.			
RP	TISSUE=Liver;			
RC	MEDLINE=96102059: Pubmed=8530402;			
RA	Rippberger J.A., Fritz S., Richter K., Hocke G.M., Lottspeich F.,			
RA	Rey G.H.;			
RT	"Transcription factors Stat3 and Stat5b are present in rat liver			
RT	nuclei late in an acute phase response and bind interleukin-6			
RT	response elements*";			

RL J. Biol. Chem. 270:29998-30006(1995).
 CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE INTERLEUKIN-6
 CC (IL-6)-RESPONSE ELEMENTS IDENTIFIED IN THE PROMOTERS OF VARIOUS
 CC ACUTE-PHASE PROTEIN GENES.
 CC -1- PATHWAY: INVOLVED IN THE Gp130-MEDIATED SIGNALING PATHWAY.
 CC -1- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
 CC MEMBER (AT LEAST START1) (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
 CC RESPONSE TO PHOSPHORYLATION (BY SIMILARITY).
 CC -1- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IL-6, IL-11, CNTF.
 CC LIF, CSF-1, EGF, PDGF, IFN-ALPHA AND OSM. SERINE PHOSPHORYLATION
 CC IS IMPORTANT FOR THE FORMATION OF STABLE DNA-BINDING STAT3
 CC HOMODIMERS AND MAXIMAL TRANSCRIPTIONAL ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X91810; CAA62920.1; -
 DR HSSP: P42224; 1BF5.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF01017; STAT; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS50001; SH2; 1.
 DR Transcription regulation; DNA-binding; Nuclear protein;
 KW Phosphorylation; SH2 domain.
 FT DOMAIN 580 670 SH2.
 FT MOD_RES 705 705 PHOSPHORYLATION (BY JAKS) (BY
 FT MOD_RES 705 705 SIMILARITY).
 FT MOD_RES 727 727 PHOSPHORYLATION (BY SIMILARITY).
 FT SEQUENCE 770 AA: 88039 MW: D74A0C769454ED CRC64;

Query Match 59.6%; Score 459; DB 1; Length 770;

Best local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
 Matches 459; Conservative 0;

312 LFRNLKSAFVVEROPCMPPHPRPLVYIKGVQFTTKVRLVLFEPRLNTYOLKIKVCDKD 371
 312 LFRNLKSAFVVEROPCMPPHPRPLVYIKGVQFTTKVRLVLFEPRLNTYOLKIKVCDKD 371
 312 LFRNLKSAFVVEROPCMPPHPRPLVYIKGVQFTTKVRLVLFEPRLNTYOLKIKVCDKD 371
 372 SGDVVALRGSRKFNILGTNTKVMNMEESNNGSLSAERKHLTLREORCGNGRANCASLI 431
 372 SGDVVALRGSRKFNILGTNTKVMNMEESNNGSLSAERKHLTLREORCGNGRANCASLI 431
 372 SGDVVALRGSRKFNILGTNTKVMNMEESNNGSLSAERKHLTLREORCGNGRANCASLI 431
 432 VPEELLLTFEPFVYHOGKIDLETHSLPVVVISNCOMNANASTLWMLNPNPNVN 491
 432 VPEELLLTFEPFVYHOGKIDLETHSLPVVVISNCOMNANASTLWMLNPNPNVN 491
 432 VPEELLLTFEPFVYHOGKIDLETHSLPVVVISNCOMNANASTLWMLNPNPNVN 491
 492 FETKPPIGTMDVAVYLSWQSSSTTKRGLSIEDLTLLAEKLLPGVYVSCQITMAFK 551
 492 FETKPPIGTMDVAVYLSWQSSSTTKRGLSIEDLTLLAEKLLPGVYVSCQITMAFK 551
 492 FETKPPIGTMDVAVYLSWQSSSTTKRGLSIEDLTLLAEKLLPGVYVSCQITMAFK 551
 492 FETKPPIGTMDVAVYLSWQSSSTTKRGLSIEDLTLLAEKLLPGVYVSCQITMAFK 551
 552 ENNAGGFSFVWLDNIIDLVKRYIIALNMGYIMGFISKEBERAILSTPFGTFLLRFS 611
 552 ENNAGGFSFVWLDNIIDLVKRYIIALNMGYIMGFISKEBERAILSTPFGTFLLRFS 611
 552 ENNAGGFSFVWLDNIIDLVKRYIIALNMGYIMGFISKEBERAILSTPFGTFLLRFS 611
 612 ESSKEGCVFTWVERKDISGTOIOSVEPTTKOOLNNMSPFEIIMGYKIMATNIVSPLV 671
 612 ESSKEGCVFTWVERKDISGTOIOSVEPTTKOOLNNMSPFEIIMGYKIMATNIVSPLV 671
 612 ESSKEGCVFTWVERKDISGTOIOSVEPTTKOOLNNMSPFEIIMGYKIMATNIVSPLV 671
 672 VLYPDIPEAEAGKCYCRPESQEHPEADPGSABAYLKTGKFCVTPPTCSNTIDLPMSRPL 731
 672 VLYPDIPEAEAGKCYCRPESQEHPEADPGSABAYLKTGKFCVTPPTCSNTIDLPMSRPL 731
 672 VLYPDIPEAEAGKCYCRPESQEHPEADPGSABAYLKTGKFCVTPPTCSNTIDLPMSRPL 731

QY 732 DSLMOPNNMGEAGPSAGCFESTLTFPMDLTSSECATSPM 770
 Db 732 DSLMOPNNMGEAGPSAGCFESTLTFPMDLTSSECATSPM 770

RESULT 3
 ID STAT3_HUMAN STANDARD; PRT; 770 AA.
 AC P40763;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 3 (ACUTE-PHASE
 DE RESPONSE FACTOR).
 GN STAT3 OR APRF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=94208062; PubMed=7512451;
 RA Akira S., Nishio Y., Inoue M., Wang X.-J., Wei S., Matsusaka T.,
 RA Yoshida K., Sudo T., Naruto M., Kishimoto T.;
 RT "Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-
 RT related transcription factor involved in the gp130-mediated signaling
 RT pathway";
 RL Cell 77:63-71(1994).
 RN [2]
 RP PHOSPHORYLATION ON SERINE.
 RX MEDLINE=95215843; PubMed=7701321;
 RA Zhang X., Fienis J., Li H.-C., Schindler C., Chen-Kiang S.;
 RT "Requirement for serine phosphorylation for formation of STAT-promoter
 RT complexes";
 RL Science 267:1990-1994(1995).
 CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE INTERLEUKIN-6
 CC (IL-6)-RESPONSIVE ELEMENTS IDENTIFIED IN THE PROMOTERS OF VARIOUS
 CC ACUTE-PHASE PROTEIN GENES.
 CC -1- PATHWAY: INVOLVED IN THE Gp130-MEDIATED SIGNALING PATHWAY.
 CC -1- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
 CC MEMBER (AT LEAST START1).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
 CC RESPONSE TO PHOSPHORYLATION.
 CC -1- TISSUE SPECIFICITY: HEART, BRAIN, PLACENTA, LUNG, LIVER, SKELETAL
 CC MUSCLE, KIDNEY, AND PANCREAS.
 CC -1- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IL-6, IL-11, CNTF.
 CC LIF, CSF-1, EGF, PDGF, IFN-ALPHA AND OSM. SERINE PHOSPHORYLATION
 CC IS IMPORTANT FOR THE FORMATION OF STABLE DNA-BINDING STAT3
 CC HOMODIMERS AND MAXIMAL TRANSCRIPTIONAL ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL: L29277; AA56374.1; -
 DR HSSP: P42224; 1BF5.
 DR TRANSFAC: T01493; -
 DR MIM: 102582; -
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF01017; STAT; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS50001; SH2; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein;
 KW Phosphorylation; SH2 domain.

```

FT DOMAIN 580 670 SH2.
FT MOD_RES 705 705 PHOSPHORYLATION (BY JAKS) (BY
FT MOD_RES 727 727 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 727 727 PHOSPHORYLATION (BY SIMILARITY).
SO SEQUENCE 770 AA: 88052 MW: A3DCEBF15B3B5360 CRC64:

Query Match 28.7%; Score 221; DB 1; Length 770;
Best Local Similarity 100.0%; Pred. No. 5,1e-220;
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 FPMELRQFLAPWIESQOMAYAAKESHATLVFNHLGELIDQYSRFLOESNVLQYHMLR 85
DB 26 FPMELRQFLAPWIESQOMAYAAKESHATLVFNHLGELIDQYSRFLOESNVLQYHMLR 85
QY 86 IKOFLQSRYLEKPMELIARIARCLMEESRLQTAATAAGCGAANHTPAAVTEKQOMLE 145
DB 86 IKOFLQSRYLEKPMELIARIARCLMEESRLQTAATAAGCGAANHTPAAVTEKQOMLE 145
QY 146 QHLDVKKRVQDLEQKKKKYVENLQDDFDFNYKTKLSGDMQDLNGNOSTRQKMOLEQ 205
DB 146 QHLDVKKRVQDLEQKKKKYVENLQDDFDFNYKTKLSGDMQDLNGNOSTRQKMOLEQ 205
QY 206 MTLALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADMKRR 246
DB 206 MTLALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADMKRR 246

RESULT 4
STAT_MOUSE STANDARD; PRT; 749 AA.
ID STAT_MOUSE
AC P42225;
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 1.
GN STAT1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
RX MEDLINE=94255416; PubMed=7545930;
RA Zhong Z., Wen Z., Darnell J.E. Jr.;
RT "Stat3 and Stat4: members of the family of signal transducers and
RT activators of transcription."
RL Proc. Natl. Acad. Sci. U.S.A. 91:4806-4810(1994).
CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE IFN-STIMULATED
CC RESPONSE ELEMENT (ISRE) AND TO THE GAS ELEMENT. THIS MULTIPROTEIN
CC TRANSCRIPTION FACTOR IS TERMED ISGF3.
CC -1- SUBUNIT: IN RESPONSE TO IFN ALPHA/BETA, THREE SUBUNITS (STAT1-
CC ALPHA, STAT1-BETA, STAT2) OF ISGF3 BECOME PHOSPHORYLATED ON
CC TYROSINE, MIGRATE INTO THE NUCLEUS, AND ASSEMBLE INTO A COMPLEX
CC TOGETHER WITH ISGF3 GAMMA (P48), A DNA-BINDING PROTEIN THAT
CC SPECIFICALLY BINDS TO THE IFN-STIMULATED RESPONSE ELEMENT. IN
CC RESPONSE TO IFN GAMMA, STAT1 FORMS HOMODIMERS, THAT ALSO
CC TRANSLocate INTO THE NUCLEUS TO ACTIVATE IFN GAMMA-RESPONSIVE
CC GENES. INTERACTS WITH NMI (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLocATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION.
CC -1- INDUCTION: BY IFN AND EGF.
CC -1- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IFN-GAMMA, IFN-ALPHA,
CC PDGF, AND EGF. SERINE PHOSPHORYLATION IS ALSO REQUIRED FOR MAXIMAL
CC TRANSCRIPTIONAL ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U06924; AAA19454.1; -.
DR HSSP: P42224; 1BF5.
DR TRANSFAC: T01575; -.
DR MGD: MGI:103063; Stat1.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein; SH2 domain;
KW Phosphorylation.
FT DOMAIN 573 670
FT MOD_RES 701 701 SH2.
FT MOD_RES 701 701 PHOSPHORYLATION (BY JAKS) (BY
FT MOD_RES 701 701 SIMILARITY).
SO SEQUENCE 727 AA: 87197 MW: 249019952BE65F1 CRC64:

Query Match 1.6%; Score 12; DB 1; Length 749;
Best Local Similarity 100.0%; Pred. No. 0.00066;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 585 IMGFISKEKERRA 596
DB 578 IMGFISKEKERRA 589

RESULT 5
STAT_HUMAN STANDARD; PRT; 750 AA.
ID STAT_HUMAN
AC P42224;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 1-ALPHA/BETA
DE (TRANSCRIPTION FACTOR ISGF-3 COMPONENTS P91/P84).
GN STAT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 514-524; 654-660 AND 667-672.
RX MEDLINE=92366557; PubMed=1502203;
RA Schindler C., Fu X.-Y., Improbata T., Aebersold R., Darnell J.E. Jr.;
RT "Proteins of transcription factor ISGF-3: one gene encodes the 91-and
RT 84-kDa ISGF-3 proteins that are activated by interferon alpha."
RL Proc. Natl. Acad. Sci. U.S.A. 89:7836-7839(1992).
RN [2]
RP PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=95192056; PubMed=7885841;
RA Yan R., Qureshi S., Zhong Z., Wen Z., Darnell J.E. Jr.;
RT "The genomic structure of the STAT genes: multiple exons in
RT coincident sites in Stat1 and Stat2."
RL Nucleic Acids Res. 23:459-463(1995).
RN [3]
RP PHOSPHORYLATION OF TYR-701.
RX MEDLINE=95386533; PubMed=7657660;
RA Quella F.W., Thierfelder W., Wiltuhn B.A., Tang B., Cohen S.,
RA Igle J.N.;
RT "Phosphorylation and activation of the DNA binding activity of
RT purified Stat1 by the Janus protein-tyrosine kinases and the
RT epidermal growth factor receptor."
RN J. Biol. Chem. 270:20775-20780(1995).
RN [4]
RP PHOSPHORYLATION OF SER-727, AND MUTAGENESIS.
RX MEDLINE=95354205; PubMed=7543024;
RA Wen Z., Zhong Z., Darnell J.E. Jr.;

```

RT "Maximal activation of transcription by Stat1 and Stat3 requires both
 RT tyrosine and serine phosphorylation."
 RL Cell 82:241-250(1995).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 136-710.
 RX MEDLINE=98292180; PubMed=9630226;
 RA Chen X., Vinkemeier U., Zhao Y., Jeruzalmi D., Darnell J.E. Jr.,
 KA Kutlany J.,
 RT "Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to
 RT DNA."
 RL Cell 93:827-839(1998).
 CC -I- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE IFN-STIMULATED
 CC RESPONSE ELEMENT (ISRE) AND TO THE GAS ELEMENT. THIS MULTIPROTEIN
 CC TRANSCRIPTION FACTOR IS TERMED ISGF3.
 CC -I- PATHWAY: IFN-SIGNALING PATHWAY.
 CC -I- SUBUNIT: IN RESPONSE TO IFN ALPHA/BETA, THREE SUBUNITS (STAT1-
 CC ALPHA, STAT1-BETA, STAT2) OF ISGF3, BECOME PHOSPHORYLATED ON
 CC TYROSINE, MIGRATE INTO THE NUCLEUS, AND ASSEMBLE INTO A COMPLEX
 CC TOGETHER WITH ISGF3 GAMMA (P48), A DNA-BINDING PROTEIN THAT
 CC SPECIFICALLY BINDS TO THE IFN-STIMULATED RESPONSE ELEMENT. IN
 CC RESPONSE TO IFN GAMMA, STAT1 FORMS HOMODIMERS, THAT ALSO
 CC TRANSLOCATE INTO THE NUCLEUS TO ACTIVATE IFN GAMMA-RESPONSIVE
 CC GENES. INTERACTS WITH NFI.
 CC -I- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
 CC RESPONSE TO PHOSPHORYLATION.
 CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS, ALPHA/P91 (SHOWN HERE) AND
 CC BETA/P84; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -I- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IFN-GAMMA, IFN-ALPHA,
 CC PDGF, AND EGF. SERINE PHOSPHORYLATION IS ALSO REQUIRED FOR MAXIMAL
 CC TRANSCRIPTIONAL ACTIVITY (LACKING IN BETA FORM).
 CC -I- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
 CC -I- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M97935; AAB64012.1; -;
 DR EMBL: M97936; -; NOT_ANNOTATED_CDS.
 DR EMBL: U18662; -; NOT_ANNOTATED_CDS.
 DR EMBL: U18663; -; NOT_ANNOTATED_CDS.
 DR EMBL: U18664; -; NOT_ANNOTATED_CDS.
 DR EMBL: U18665; -; NOT_ANNOTATED_CDS.
 DR EMBL: U18666; -; NOT_ANNOTATED_CDS.
 DR EMBL: U18667; -; NOT_ANNOTATED_CDS.
 DR EMBL: U18668; -; NOT_ANNOTATED_CDS.
 DR EMBL: U18669; -; NOT_ANNOTATED_CDS.
 DR EMBL: U18670; -; NOT_ANNOTATED_CDS.
 DR PDB: 1BF5; 12-AUG-98.
 DR TRANSFAC: T01492; -;
 DR MIM: 600555; -;
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF00017; SH2; 1.
 DR SMART: PF01017; STAT; 1.
 DR SMART: SMO0252; SH2; 1.
 DR PROSITE: PS50001; SH2; 1.
 DR Transcription regulation: DNA-binding; Nuclear protein;
 DR Phosphorylation; SH2 domain; Alternative splicing; 3D-structure.
 KM DOMAIN 573 670
 FT MOD_RES 701 701 PHOSPHORYLATION (BY JAKS).
 FT MOD_RES 727 727 PHOSPHORYLATION.
 FT VARSPIC 713 750 MISSING (IN ISOFORM BETA).
 FT MUTAGEN 727 727 S->A: DECREASED TRANSCRIPTIONAL
 FT ACTIVATION.
 FT CONFLICT 261 271 MISSING (IN M97936).
 FT SEQUENCE 750 AA; 87334 MW; 054A81352364BA6 CRC64;

Query Match 1.6%; Score 12; DB 1; Length 750;
 Best Local Similarity 100.0%; Pred. No. 0.00066;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 585 IMGFISKERERA 596
 DB 578 IMGFISKERERA 589

RESULT 6
 STA4_HUMAN
 ID STA4_HUMAN STANDARD; PRI: 748 AA.
 AC 014765;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 4.
 GN STA4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Xu X., Sun Y.L., Hoey T.;
 RT "The STAT amino-terminal domain mediates cooperative DNA binding
 RT and confers selective sequence recognition."
 RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
 CC -I- FUNCTION: CARRIES OUT A DUAL FUNCTION: SIGNAL TRANSDUCTION AND
 CC ACTIVATION OF TRANSCRIPTION.
 CC -I- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
 CC MEMBER (BY SIMILARITY).
 CC -I- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
 CC RESPONSE TO PHOSPHORYLATION.
 CC -I- PTM: TYROSINE PHOSPHORYLATED. SERINE PHOSPHORYLATION IS ALSO
 CC REQUIRED FOR MAXIMAL TRANSCRIPTIONAL ACTIVITY (BY SIMILARITY).
 CC -I- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
 CC -I- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L78440; AAB05605.1; -;
 DR MIM: 600558; -;
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF00017; SH2; 1.
 DR SMART: PF01017; STAT; 1.
 DR SMART: SMO0252; SH2; 1.
 DR PROSITE: PS50001; SH2; 1.
 DR Transcription regulation: DNA-binding; Nuclear protein;
 KM Phosphorylation; SH2 domain.
 FT DOMAIN 569 664
 FT MOD_RES 693 693 PHOSPHORYLATION (BY JAKS) (BY
 FT SIMILARITY).
 FT MOD_RES 721 721 PHOSPHORYLATION (BY SIMILARITY).
 FT SEQUENCE 748 AA; 85940 MW; 11E43803A9AFAFFA CRC64;

Query Match 1.4%; Score 11; DB 1; Length 748;
 Best Local Similarity 100.0%; Pred. No. 0.0071;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 603 PGFPLRSES 613
 DB 592 PGFPLRSES 602

```

RESULT 7
STA2_MOUSE STANDARD: PRT: 749 AA.
ID STA2_MOUSE
AC P42228:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 4.
GN STA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
RA MEDLINE=94255416; PubMed=7545930;
RA Zhong Z., Wen Z., Darnell J.E. Jr.;
RT "Stat3 and Stat4: members of the family of signal transducers and
RT activators of transcription."
RT Proc. Natl. Acad. Sci. U.S.A. 91:4806-4810(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Testis;
RX MEDLINE=94277038; PubMed=8007943;
RX Yamamoto K., Quelle F.W., Thierfelder W.E., Kreider B.L.,
RX Gilbert D.J., Jenkins N.A., Copeland N.G., Silvennoinen O.,
RX Thle J.N.;
RT "Stat4, a novel gamma interferon activation site-binding protein
RT expressed in early myeloid differentiation."
RT Mol. Cell. Biol. 14:4342-4349(1994).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.45 ANGSTROMS) OF 1-123.
RX MEDLINE=98128033; PubMed=9461439;
RX Vinkemeier U., Moarefi I., Darnell J.E. Jr., Kuriyan J.;
RT "Structure of the amino-terminal protein interaction domain of
RT Stat-4."
RL Science 279:1048-1052(1998).
CC -1- FUNCTION: CARRIES OUT A DUAL FUNCTION: SIGNAL TRANSDUCTION AND
CC ACTIVATION OF TRANSCRIPTION.
CC -1- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
CC MEMBER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION.
CC -1- PTM: TYROSINE PHOSPHORYLATED. SERINE PHOSPHORYLATION IS ALSO
CC REQUIRED FOR MAXIMAL TRANSCRIPTIONAL ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: U06923; AAA19453.1; -
CC EMBL: U09351; AAA19692.1; -
CC PDB: 1BGF; 16-SEP-98.
CC TRANSFAC: T01576; -
CC MGD: MGI:103062; Stat4.
CC InterPro: IPR000980; SH2.
CC InterPro: IPR001217; STAT.
CC Pfam: PF00017; SH2_1.
CC Pfam: PF01017; STAT_1.
CC SMART: SM00522; SH2; 1.
CC PROSITE: PS50001; SH2; 1.
CC Transcription regulation: DNA-binding; Nuclear protein;
CC Phosphorylation: SH2 domain; 3D-structure.
CC KJ DOMAIN 570 665
CC FT MOD_RES 694 694
CC SIMILARITY).

```

```

FT MOD_RES 722 722 PHOSPHORYLATION (BY SIMILARITY).
FT CONFLICT 199 199 MISSING (IN REF. 2).
FT CONFLICT 638 638 A -> P (IN REF. 2).
SQ SEQUENCE 749 AA: 85940 MW: A88B837E49CFCEBC CRC64:

Query Match 1.4%; Score 11; DB 1; Length 749;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 603 PGFFLLRFSES 613
Db 593 PGFFLLRFSES 603

RESULT 8
STA2_HUMAN STANDARD: PRT: 851 AA.
ID STA2_HUMAN
AC P52630; Q16430; Q16431;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 2 (P113).
GN STAT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95192056; PubMed=7885841;
RX Yan R., Qureshi S., Zhong Z., Wen Z., Darnell J.E. Jr.;
RT "The genomic structure of the STAT genes: multiple exons in
RT coincident sites in Stat1 and Stat2."
RL Nucleic Acids Res. 23:459-463(1995).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=9236558; PubMed=1502204;
RX Fu X.-Y., Schindler C., Impirota T., Aebersold R., Darnell J.E. Jr.;
RT "The proteins of ISGF-3, the interferon alpha-induced transcriptional
RT activator, define a gene family involved in signal transduction."
RL Proc. Natl. Acad. Sci. U.S.A. 89:7840-7843(1992).
RN [3]
RP ALTERNATIVE SPLICING.
RX MEDLINE=96176320; PubMed=8601453;
RX Sugiyama T., Nishio Y., Kishimoto T., Akira S.;
RT "Identification of alternative splicing form of Stat2."
RL FEBS Lett. 381:191-194(1996).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=9020188;
RX Bluyssen H.A., Levy D.E.;
RT "Stat2 is a transcriptional activator that requires sequence-specific
RT contacts provided by stat1 and p48 for stable interaction with DNA."
RL J. Biol. Chem. 272:4600-4605(1997).
CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE IFN-STIMULATED
CC RESPONSE ELEMENT (ISRE) AND TO THE GAS ELEMENT. THIS MULTIPROTEIN
CC TRANSCRIPTION FACTOR IS TERMED ISGF3.
CC -1- SUBUNIT: IN RESPONSE TO IFN ALPHA/BETA, THREE SUBUNITS (STAT1-
CC ALPHA, STAT1-BETA, STAT2) OF ISGF3, BECOME PHOSPHORYLATED ON
CC TYROSINE, MIGRATE INTO THE NUCLEUS, AND ASSEMBLE INTO A COMPLEX
CC TOGETHER WITH ISGF3 GAMMA (P48). A DNA-BINDING PROTEIN THAT
CC SPECIFICALLY BINDS TO THE IFN-STIMULATED RESPONSE ELEMENT.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IFN-ALPHA.
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U18671: AAA8760.1: -
DR EMBL: M97934: -; NOT_ANNOTATED_CDS.
DR EMBL: S81491: AAB36226.1: -
DR EMBL: S81491: AAB36227.1: ALT_SPO.
DR HSSP: P42224: 1BF5.
DR MIM: 600556: -
DR InterPro: IPR000980: SH2.
DR InterPro: IPR001217: STAT.
DR Pfam: PF00017: SH2; 1.
DR Pfam: PF01017: STAT; 1.
DR SMART: SM00252: SH2; 1.
DR PROSITE: PS50001: SH2; 1.
KM Transcription regulation; DNA-binding; Nuclear protein;
KW Phosphorylation; SH2 domain; Alternative splicing.
FT DOMAIN 572 667 SH2.
FT MOD_RES 690 690 PHOSPHORYLATION (BY JAKS) (BY
SIMILARITY).
FT VARSPIC 621 652 DKVLIVSOPYREVLQSLPLTEIRHYQLLT -> GQLLC
FT PAIPASPECLAPLPLCLPLASLALN (IN SHORT
ISOFORM).
FT VARSPIC 653 851 MISSING (IN SHORT ISOFORM).
FT SEQUENCE 851 AA: 97916 MW: 6474674CB7A3215 CRC64;

Query Match 1.4%; Score 11; DB 1; Length 851;
Best Local Similarity 100.0%; Pred. No. 0.0079;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 604 GTFILRFSESS 614
DB 596 GTFILRFSESS 606
IIIIIIIIII

RESULT 9
STA2_PIG STANDARD: PRT: 864 AA.
AC 002799;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 2.
GN STAT2
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Ito Y., Mikawa S., Kobayashi E., Wada Y., Minezawa M.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE IFN-STIMULATED
CC RESPONSE ELEMENT (ISRE) AND TO THE GAS ELEMENT. THIS MULTIPROTEIN
CC TRANSCRIPTION FACTOR IS TERMED ISGF3 (BY SIMILARITY).
CC -1- SUBUNIT: IN RESPONSE TO IFN ALPHA/BETA, THREE SUBUNITS (STAT1-
CC ALPHA, STAT1-BETA, STAT2) OF ISGF3, BECOME PHOSPHORYLATED ON
CC TYROSINE, MIGRATE INTO THE NUCLEUS, AND ASSEMBLE INTO A COMPLEX
CC TOGETHER WITH ISGF3 GAMMA (P48), A DNA-BINDING PROTEIN THAT
CC SPECIFICALLY BINDS TO THE IFN-STIMULATED RESPONSE ELEMENT (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION (BY SIMILARITY).
CC -1- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IFN-ALPHA (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.

CC This SWISS entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AB004061: BAA20332.1: -
DR InterPro: IPR000980: SH2.
DR InterPro: IPR001217: STAT.
DR Pfam: PF00017: SH2; 1.
DR Pfam: PF01017: STAT; 1.
DR SMART: SM00252: SH2; 1.
DR PROSITE: PS50001: SH2; 1.
KM Transcription regulation; DNA-binding; Nuclear protein;
KW Phosphorylation; SH2 domain.
FT DOMAIN 572 667 SH2.
FT MOD_RES 690 690 PHOSPHORYLATION (BY JAKS) (BY
SIMILARITY).
FT SEQUENCE 864 AA: 99060 MW: 2F8CC95569FB5B20 CRC64;

Query Match 1.2%; Score 9; DB 1; Length 864;
Best Local Similarity 100.0%; Pred. No. 0.094;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 604 GTFILRFSE 612
DB 596 GTFILRFSE 604
IIIIIIIIII

RESULT 10
STA2_MOUSE STANDARD: PRT: 923 AA.
AC Q9WV12: Q64189; Q64250; Q64188;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 2.
GN STAT2
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1;
RA Paulson N.S., Mui A., Levy D.E.;
RT "Molecular cloning and characterization of murine Stat2."
RL Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE OF 595-658 FROM N.A. (ISOFORMS A AND B/C).
RX MEDLINE=96176320; Pubmed=8601453;
RA Sugiyama T., Nishio Y., Kishimoto T., Akira S.;
RL "Identification of alternative splicing form of Stat2."
CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE IFN-STIMULATED
CC RESPONSE ELEMENT (ISRE) AND TO THE GAS ELEMENT. THIS MULTIPROTEIN
CC TRANSCRIPTION FACTOR IS TERMED ISGF3.
CC -1- SUBUNIT: IN RESPONSE TO IFN ALPHA/BETA, THREE SUBUNITS (STAT1-
CC ALPHA, STAT1-BETA, STAT2) OF ISGF3, BECOME PHOSPHORYLATED ON
CC TYROSINE, MIGRATE INTO THE NUCLEUS, AND ASSEMBLE INTO A COMPLEX
CC TOGETHER WITH ISGF3 GAMMA (P48), A DNA-BINDING PROTEIN THAT
CC SPECIFICALLY BINDS TO THE IFN-STIMULATED RESPONSE ELEMENT (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A (SHOWN HERE) AND B/C; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: FOUND IN THE BRAIN, LUNG, HEART, SPLEEN,
CC LIVER, KIDNEY, MUSCLE, AND THE TESTIS.
CC -1- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IFN-ALPHA (BY

```

CC      1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC      1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF088862; AAD38329.1; -
DR      EMBL; S81342; AAB36228.2; -
DR      EMBL; S81342; AAB36231.1; -
DR      EMBL; S81342; AAB36230.1; ALT_SEQ.
DR      HSSP; P42224; 1BF5.
DR      MCD; MGI:103039; Stat2.
DR      InterPro: IPR000980; SH2.
DR      InterPro: IPR001217; STAT.
DR      Pfam; PF00017; SH2; 1.
DR      Pfam; PF01017; STAT; 1.
DR      SMART; SM00252; SH2; 1.
DR      PROSITE; PS50001; SH2; 1.
KW      Transcription regulation; DNA-binding; Nuclear protein;
KW      phosphorylation; SH2 domain; Alternative splicing.
FT      DOMAIN 571 666
FT      MOD_RES 689 689
FT      PHOSPHORYLATION (BY JAKS) (BY
FT      SIMILARITY).
FT      VARSPIC 620 643
FT      HKEVYSQPYTKELQSLPLEI -> GQHPPYHSCSL
FT      SARHPITRLP (IN SHORT ISOFORM).
FT      MISSING (IN SHORT ISOFORM).
FT      T -> A (IN REF. 2).
FT      H -> D (IN REF. 2).
FT      CONFLICT 620 620
SO      SEQUENCE 923 AA: 105416 MW: D50BB54C35B0774 CRC64;

Query Match 1.2%; Score 9; DB 1; Length 923;
Best Local Similarity 100.0%; Pred. No. 0.99;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 604 GTFLLRFSE 612
    |||||
Db 595 GTFLLRFSE 603

RESULT 11
HDNO_ARTOX STANDARD: PRT: 458 AA.
AC P08159;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 6-HYDROXY-D-NICOTINE OXIDASE (EC 1.5.3.6) (6-HDNO).
OS Arthrobacter oxidans.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococcales; Micrococcaceae; Arthrobacter.
OC NCBI_TaxID=1671;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=87304263; PubMed=3622516;
RA Brandesch R., Hinkkanen A.E., Mauch L., Nagursky H., Decker K.;
RT "6-Hydroxy-D-nicotine oxidase of Arthrobacter oxidans. Gene structure
RT of the flavoenzyme and its relationship to 6-hydroxy-L-nicotine
RT oxidase."
RL Eur. J. Biochem. 167:315-320(1987).
RN 12
RP REVISIONS.
RA Brandesch R.;
RX Submitted (SEP-1990) to the EMBL/GenBank/DBJ databases.
RN 13
RP MUTAGENESIS OF HIS-71.
RX MEDLINE=90033359; PubMed=2680607;

```

```

RA Mauch L., Biehler V., Brandesch R.;
RT "Site-directed mutagenesis of the FAD-binding histidine of
RT 6-hydroxy-D-nicotine oxidase. Consequences on flavinylation and
RT enzyme activity."
RL FEBS Lett. 257:86-88(1989).
RN 14
RP MUTAGENESIS OF ARG-67 AND SER-68.
RX MEDLINE=90330600; PubMed=2115879;
RA Mauch L., Biehler V., Brandesch R.;
RT "Lysine can replace arginine 67 in the mediation of covalent
RT attachment of FAD to histidine 71 of 6-hydroxy-D-nicotine oxidase."
RL J. Biol. Chem. 265:12761-12762(1990).
CC 1- CATALYTIC ACTIVITY: (D)-6-HYDROXYNICOTINE + H(2)O + O(2) =
CC 1-(6-HYDROXY)PYRID-3-YL)-4-(METHYLAMINO)BUTAN-1-ONE + H(2)O(2).
CC 1- COFACTOR: FAD.
CC 1- PATHWAY: DEGRADATION OF NICOTINE (INDUCIBLE).
CC 1- SIMILARITY: TO OTHER OXYGEN OXIDOREDUCTASES THAT COVALENTLY BIND
CC FAD.
CC -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X05999; CAA29416.1; -
DR      PIR; S00087; DE10HN.
DR      InterPro: IPR001575; Oxid_FAD_bind.
DR      Pfam; PF01565; FAD_binding_4; 1.
DR      PROSITE; PS00862; OX2_COVAL_FAD; 1.
KW      Oxidoreductase; Flavoprotein; FAD.
FT      BINDING 71 71
FT      MUTAGEN 67 67
FT      MUTAGEN 67 67
FT      MUTAGEN 68 68
FT      MUTAGEN 71 71
FT      SEQUENCE 458 AA: 48786 MW: 6783E44AD66DC841 CRC64;

Query Match 1.0%; Score 8; DB 1; Length 458;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 218 VSELAGLL 225
    |||||
Db 219 VSELAGLL 226

RESULT 12
MNEI_YEAST STANDARD: PRT: 663 AA.
AC P24720;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE MNEI PROTEIN.
GN MNEI OR MNE OR YOR350C OR O6353.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN 11
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / FY73;
RX MEDLINE=97103776; PubMed=8948102;
RA Purnelle B., Goffeau A.;
RT "Nucleotide sequence analysis of a 40 kb segment on the right arm of
RT yeast chromsome XV reveals 18 open reading frames including a new

```

```

RT pyruvate kinase and three homologues to chromosome 1 genes.*;
RL Yeast 12:1475-1481(1996).
[2]
RP SEQUENCE OF 1-219 FROM N.A.
RX MEDLINE-92158649; Pubmed-1741279;
RA Leem S.H., Ogawa H.;
RT "The MRE4 gene encodes a novel protein kinase homologue required for
RL meiotic recombination in Saccharomyces cerevisiae.";
CC Nucleic Acids Res. 20:449-457(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X95720; CAA65037.1; -
DR EMBL: 275258; CAA99678.1; -
DR EMBL: X63112; CAA44826.1; -
DR PIR: S19073; S19073.
DR PIR: S20175; S20175.
DR SGD: S0005877; MNE1.
SQ SEQUENCE 663 AA; 77229 MW; 208DD629E7034858 CRC64;

Query Match 1.08; Score 8; DB 1; Length 663;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 590 SKERERAT 597
Db 249 SKERERAT 256

RESULT 13
STAT_DROME STANDARD; PRT; 761 AA.
AC 024151: 024181: 09YDL8:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SIGNAL TRANSDUCER AND TRANSCRIPTION ACTIVATOR (MARELLE PROTEIN)
DE (D-STAT).
GN STAT92E OR MARE OR MRL OR STAT OR CG4257.
OS Drosophila melanogaster (Fruit fly)
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A. (LONG ISOFORM).
RC TISSUE=Embryo.
RX MEDLINE-96190808; Pubmed-8608595.
RA Hou X.S., Meinick M.B., Perrimon N.;
RT "Marelle acts downstream of the Drosophila HOP/JAK kinase and encodes
RL a protein similar to the mammalian STATs.";
CC Cell 84:411-419(1996).
RN [2]
RP SEQUENCE FROM N.A. (SHORT ISOFORM).
RC TISSUE=Embryo.
RX MEDLINE-96190809; Pubmed-8608596;
RA Yan R., Small S., Desplan C., Dearolf C.R., Darnell J.E. Jr.;
RT "Identification of a Stat gene that functions in Drosophila
RL development.";
CC Cell 84:421-430(1996).
RN [3]
RP SEQUENCE FROM N.A. (LONG ISOFORM).
RC STRAIN=BERKELEY.
RX MEDLINE-20196006; Pubmed-10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,

```

```

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sulton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Abdanyani A., An H.-J., Andrews-Pfennoch C., Baldwin D.,
RA Balcer R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benus P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokhov D., Botchan M.R., Bouck J., Brackstein P., Brothier P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
RA de Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Kochia S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischman W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hartley D., Harvey D., Heiman T.J., Hernandez J.R., Honick J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spter E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhen M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong M., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: SIGNAL TRANSDUCTION AND ACTIVATION OF TRANSCRIPTION.
CC PLAYS AN IMPORTANT ROLE IN THE SEGMENTAL PATTERN FORMATION IN THE
CC EARLY EMBRYO BY ACTIVATING SPECIFIC STRIPS OF PAIR RULE GENE
CC EXPRESSION. THE JAKUS KINASE-STAT PATHWAY IS CONNECTED TO
CC DROSOPHILA EARLY DEVELOPMENT. MUTANTS EXHIBIT ABERRANT EXPRESSION
CC OF THE PAIR RULE GENE EVEN-SKIPPED AT THE CELLULAR BLASTODERM
CC STAGE, LEADING TO LARVAL SEGMENTATION DEFECTS.
CC -1- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
CC MEMBER.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND
CC A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT EMBRYONIC, LARVAL,
CC PUPAL, AND ADULT STAGES, WITH SOME DECREASE IN THE LATE EMBRYONIC
CC STAGES. THE EXPRESSION IS UNIFORM IN UNFERTILIZED OR JUST
CC FERTILIZED EGGS, SUGGESTING MATERNALLY DEPOSITED mRNA. AT
CC BLASTODERM STAGE, EXPRESSION PATTERN SHOWS STRIPS, THAT ARE
CC REMINISCENT OF MANY PAIR RULE GENES PATTERN.
CC -1- PTM: TYROSINE PHOSPHORYLATED BY HOSCOCH. PHOSPHORYLATION IS
CC REQUIRED FOR DNA-BINDING ACTIVITY AND DIMERIZATION.
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC IN PARTICULAR TO MAMMALIAN STATs AND STATc.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U40070; AAC46984.1; -
DR EMBL: U46688; AAB02195.1; -
DR EMBL: AE003731; AAF55773.1; -
DR HSSP: P42224; 1BF5.

```

```

DR FlyBase; FBgn0016917; Stat92E.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PF01017; STAT; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PSS0001; SH2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein; SH2 domain;
KW Phosphorylation; Alternative splicing; Developmental protein.
FT DOMAIN 594 658 SH2.
FT MOD_RES 711 711 PHOSPHORYLATION (BY JAKS).
FT VARSPIC 699 705 MISSING (IN SHORT ISOFORM).
FT CONFLICT 105 105 S -> T (IN REF. 2).
FT CONFLICT 648 648 L -> H (IN REF. 2).
SQ SEQUENCE 761 AA; 86414 MW; BF3A622A29899161 CRC64;

Query Match 1.0%; Score 8; DB 1; Length 761;
Best Local Similarity 100.0%; Pred. No. 9.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 604 GTFLLRFS 611
DB 620 GTFLLRFS 627

RESULT 14
ST5B_HUMAN STANDARD; PRT; 786 AA.
ID ST5B_HUMAN
AC P51692;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 5B.
GN STAT5B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RX MEDLINE=96311205; PubMed=87332682;
RA Silva C.M., Lu H., Day R.N.;
RT "Characterization and cloning of STAT5 from IM-9 cells and its
RL activation by growth hormone.";
RL Mol. Endocrinol. 10:508-518(1996).
RN [2]
RX SEQUENCE FROM N.A.
RX MEDLINE=96210005; PubMed=8631883;
RA Lin J.-X., Metz J., Modi W.S., John S., Leonard W.J.;
RT "Cloning of human Stat5B. Reconstitution of interleukin-2-induced
RL Stat5A and Stat5B DNA binding activity in COS-7 cells.";
RL J. Biol. Chem. 271:10738-10744(1996).
RN [3]
RX INTERACTION WITH NMI.
RX PubMed=9989503;
RA Zhu M.-H., John S., Berg M., Leonard W.J.;
RT "Functional association of Nmi with Stat5 and Stat1 in IL-2- and
RT IFNgamma-mediated signaling.";
RL Cell 96:121-130(1999).
CC -1- FUNCTION: CARRIES OUT A DUAL FUNCTION: SIGNAL TRANSDUCTION AND
CC ACTIVATION OF TRANSCRIPTION. BINDS TO THE GAS ELEMENT AND
CC ACTIVATES PRL-INDUCED TRANSCRIPTION.
CC -1- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
CC MEMBER (BY SIMILARITY). INTERACTS WITH NMI.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION (BY SIMILARITY).
CC -1- PTM: TYROSINE PHOSPHORYLATED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U48730; AAC50485.1; --
DR EMBL; U47686; AAC50491.1; --
DR HSSP; P42224; 1BF5.
DR MIM; 604260; --
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; STAT; 1.
DR SMART; SMC_52; SH2; 1.
DR PROSITE; PSS0001; SH2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW Phosphorylation; SH2 domain.
FT DOMAIN 589 686 SH2.
FT MOD_RES 699 699 PHOSPHORYLATION (BY JAKS) (BY
FT SIMILARITY).
FT DOMAIN 717 720 REQUIRED FOR INTERACTION WITH NMI.
FT CONFLICT 230 230 A -> P (IN REF. 2).
FT CONFLICT 628 628 S -> T (IN REF. 2).
FT CONFLICT 717 717 D -> DA (IN REF. 2).
FT CONFLICT 720 720 R -> G (IN REF. 2).
SQ SEQUENCE 786 AA; 89880 MW; E4265C4C8EB824B6 CRC64;

Query Match 1.0%; Score 8; DB 1; Length 786;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 604 GTFLLRFS 611
DB 613 GTFLLRFS 620

RESULT 15
ST5B_MOUSE STANDARD; PRT; 786 AA.
ID ST5B_MOUSE
AC P42232; O60804;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 5B.
GN STAT5B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
NCBI_TaxID=10090;
RX MEDLINE=95237198; PubMed=7720707;
RA Mui A.L.-F., Makao H., O'Farrell A.-M., Harada N., Miyajima A.;
RT "Interleukin-3, granulocyte-macrophage colony stimulating factor and
RT interleukin-5 transduce signals through two STAT5 homologs.";
RL EMBO J. 14:1166-1175(1995).
RN [2]
RX SEQUENCE FROM N.A.
RX STRAIN=C57BL/6;
RX MEDLINE=96004632; PubMed=7568026;
RA Liu X., Robinson G.W., Gouilleux F., Groner B., Hennighausen L.;
RT "Cloning and expression of Stat5 and an additional homologue (Stat5b)
RT involved in prolactin signal transduction in mouse mammary tissue.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:8831-8835(1995).
CC -1- FUNCTION: CARRIES OUT A DUAL FUNCTION: SIGNAL TRANSDUCTION AND
CC ACTIVATION OF TRANSCRIPTION. BINDS TO THE GAS ELEMENT AND
CC ACTIVATES PRL-INDUCED TRANSCRIPTION.
CC -1- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
CC MEMBER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION.
CC -----

```

```
CC -I- TISSUE SPECIFICITY: IN THE VIRGIN, FOUND IN MOST TISSUES.
CC PARTICULARLY ABUNDANT IN MUSCLE TISSUE OF VIRGIN AND LACTATING
CC FEMALES, AND OF MALES.
CC -I- DEVELOPMENTAL STAGE: DETECTED BOTH IN VIRGIN MOUSE AND AFTER
CC MAMMARY GLAND INVOLUTION. THE LEVEL OF STAT5A INCREASES CONSTANTLY
CC DURING PREGNANCY, BUT DECREASES DURING LACTATION.
CC -I- PTM: TYROSINE PHOSPHORYLATED.
CC -I- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -I- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z48539; CAAB8420.1; -.
DR EMBL: U21110; AAC52282.1; -.
DR HSSP: P42224; 1BP5.
DR TRANSFAC: T00944; -.
DR MGD: MGI:103035; Stat5b.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
DR Transcription regulation: DNA-binding; Nuclear protein;
KW Phosphorylation; SH2 domain.
FT DOMAIN 589 686
FT MOD_RES 699 699
FT PHOSPHORYLATION (BY JAKS) (BY
FT SIMILARITY).
FT CONFLICT 433 433 E->G (IN REF. 2).
FT SEQUENCE 786 AA: 90002 MW: A8FE76405E41B2EF CRC64:
```

```
Query Match 1.0%; Score 8; DB 1; Length 786;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 604 GTFLLRFS 611
    |||||
Db 613 GTFLLRFS 620
```

Search completed: March 19, 2002, 15:27:00
Job time: 252 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2002, 14:00:57 ; Search time 2747.57 Seconds

(without alignments)
17226.279 Million cell updates/sec

Title: US-08-212-185-11

Perfect score: 2869
Sequence: 1 GCCGCGACCGACGCGCGG.....AATTAAAAAAAAAAAAAAAA 2869

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_in:*
18: em_in:*
19: em_om:*
20: em_om:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vl:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2869	100.0	2869	6	AR083251	AR083251 Sequence
2	2869	100.0	2869	6	187790	187790 Sequence 11
3	2864.2	99.8	2869	6	AR121061	AR121061 Sequence
4	2864.2	99.8	2869	10	MMU06922	MMU06922 Mus musculus
5	2861	99.7	2869	6	AR153599	AR153599 Sequence
6	2767.8	96.5	2964	10	BC003806	BC003806 Mus muscu
7	2711.2	94.5	2924	10	MMU07099	MMU07099 Mus muscu
8	2705.2	94.3	2924	10	MMU08378	MMU08378 Mus muscu
9	2415.6	84.2	2652	6	188783	188783 Sequence 7
10	2415.6	84.2	2652	6	188784	188784 Sequence 8
11	2415.6	84.2	2652	10	MUSABRF	L29278 Mus musculu
12	2415.6	84.2	2652	22	E10408	E10408 CDNA encodi
13	2296	80.0	2924	10	RNSTRT3PT	X91810 R.norvegicu
14	2294	80.0	2310	6	188782	188782 Sequence 6
15	2144.2	74.7	2787	6	AR120980	AR120980 Sequence
16	2144.2	74.7	2787	6	AR121650	AR121650 Sequence
17	2144.2	74.7	2787	6	188780	188780 Sequence 3
18	2144.2	74.7	2787	6	188781	188781 Sequence 4
19	2144.2	74.7	2787	9	HUMAPRF	L29277 Homo sapien
20	2144.2	74.7	2787	22	E10407	E10407 CDNA encodi
21	2071.8	72.2	2663	9	BC000627	BC000627 Homo sapi
22	2026.8	70.6	2310	6	188779	188779 Sequence 2
23	2012.6	70.1	2344	6	A98087	A98087 Sequence 1
24	2012.6	70.1	2344	6	A98229	A98229 Sequence 1
25	1993	69.5	2633	5	HSA012463	A1012463 Homo sapi
26	1528.8	53.3	2653	5	AB017701	AB017701 Xenopus l
27	1389.8	48.4	3071	5	DRE5693	AJ005693 Dario rer
28	1383	48.2	2304	5	OMU60333	OMU60333 Oncorhynch
29	698.6	24.3	2643	10	BC004808	BC004808 Mus muscu
30	682	23.8	2277	6	AR083249	AR083249 Sequence
31	682	23.8	2277	6	187788	187788 Sequence 7
32	682	23.8	2277	10	MMU06924	MMU06924 Mus musculu
33	676	23.6	2260	10	AF205604	AF205604 Rattus no
34	672.6	23.4	2607	6	AR083248	AR083248 Sequence
35	672.6	23.4	2607	6	187787	187787 Sequence 5
36	672.6	23.4	2607	9	HUMISGF3B	M97936 Human trans
37	672.6	23.4	2638	9	BC002704	BC002704 Homo sapi
38	672.6	23.4	3943	6	AR083247	AR083247 Sequence
39	672.6	23.4	3943	6	187786	187786 Sequence 3
40	672.6	23.4	4003	6	AR121648	AR121648 Sequence
41	672.6	23.4	4003	6	HUMISGF3A	M97935 Homo sapien
42	666.2	23.2	3943	6	AR153600	AR153600 Sequence
43	661.4	23.1	2265	5	OMU60331	OMU60331 Oncorhynch
44	658.8	23.0	2265	5	OMU60332	OMU60332 Oncorhynch
45	580.4	20.2	2533	10	MMU06923	MMU06923 Mus musculu

ALIGNMENTS

RESULT 1
LOCUS AR083251 2869 bp DNA
DEFINITION Sequence 11 from patent US 5976835.
ACCESSION AR083251
VERSION AR083251.1 GI:10010041
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2869)
AUTHORS Darnell,J.E., Jr., Schindler,C.W., Fu,X., Wen,Z., and Zhong,Z.
TITLE Nucleic acids encoding receptor recognition factor Stat1.alpha. and
JOURNAL Patent: US 5976835-A 11 02-NOV-1999;
FEATURES Location/Qualifiers
source 1..2869
BASE COUNT 755 a 743 c 772 g 599 t
ORIGIN
/organism="unknown"

Query Match	100.0%;	Score 2869;	DB 6;	Length 2869;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2869; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Oy	1	GGCGGACACAGCCGAGGCGCGGACAGTGGGGATCTACGCCCCGAGACAGTGGACACCCCTGACT	60
Oy	1	GGCGGACACAGCCGAGGCGCGGACAGTGGGGATCTACGCCCCGAGACAGTGGACACCCCTGACT	60
Db	1	GGCGGACACAGCCGAGGCGCGGACAGTGGGGATCTACGCCCCGAGACAGTGGACACCCCTGACT	60
Oy	61	GCACGAGGATGGCTCAGTGGAAACCAAGTCCAGCAAGCTGGACACACGCTACCTGAAGACG	120
Oy	61	GCACGAGGATGGCTCAGTGGAAACCAAGTCCAGCAAGCTGGACACACGCTACCTGAAGACG	120
Db	61	GCACGAGGATGGCTCAGTGGAAACCAAGTCCAGCAAGCTGGACACACGCTACCTGAAGACG	120
Oy	121	TGCACCAAGCTGTACAGCAGACGTTCCCATATGAGCTCGGCGAGTCTCTGGACACTTGA	180
Oy	121	TGCACCAAGCTGTACAGCAGACGTTCCCATATGAGCTCGGCGAGTCTCTGGACACTTGA	180
Db	121	TGCACCAAGCTGTACAGCAGACGTTCCCATATGAGCTCGGCGAGTCTCTGGACACTTGA	180
Oy	181	TTGAGAGTCAAGACTGGGCAATATGCAAGCCAGCAAGAGTCAATGCTGCTGTTTC	240
Db	181	TTGAGAGTCAAGACTGGGCAATATGCAAGCCAGCAAGAGTCAATGCTGCTGTTTC	240
Oy	241	ATAATCTCTGGGTGAATATGACACAGCATATAGCCGATTCCTCGAAGAGTCCAAATCTCC	300
Oy	241	ATAATCTCTGGGTGAATATGACACAGCATATAGCCGATTCCTCGAAGAGTCCAAATCTCC	300
Db	241	ATAATCTCTGGGTGAATATGACACAGCATATAGCCGATTCCTCGAAGAGTCCAAATCTCC	300
Oy	301	TCTATCAGCAACACCTTTCGAAGATCAGCAAGTCTTCTGACAGCAGATATCTTAGAACG	360
Db	301	TCTATCAGCAACACCTTTCGAAGATCAGCAAGTCTTCTGACAGCAGATATCTTAGAACG	360
Oy	361	CAATGGAATTTGGCCGGATGTGGCCCGATGGCTGTGGAAAGAGTCTGCTCCACGA	420
Oy	361	CAATGGAATTTGGCCGGATGTGGCCCGATGGCTGTGGAAAGAGTCTGCTCCACGA	420
Db	361	CAATGGAATTTGGCCGGATGTGGCCCGATGGCTGTGGAAAGAGTCTGCTCCACGA	420
Oy	421	CGGACGCCAGCGACGCCACCAAGGGGGCCAGGGCAACCAACCAAGCCGCGCTACTGA	480
Oy	421	CGGACGCCAGCGACGCCACCAAGGGGGCCAGGGCAACCAACCAAGCCGCGCTACTGA	480
Db	421	CGGACGCCAGCGACGCCACCAAGGGGGCCAGGGCAACCAACCAAGCCGCGCTACTGA	480
Oy	481	CAGAGAAAGCAGAGATTTTGAGACACACATCTTCAGAGATGTCCGGAAGGAGATGCAAGATC	540
Oy	481	CAGAGAAAGCAGAGATTTTGAGACACACATCTTCAGAGATGTCCGGAAGGAGATGCAAGATC	540
Db	481	CAGAGAAAGCAGAGATTTTGAGACACACATCTTCAGAGATGTCCGGAAGGAGATGCAAGATC	540
Oy	541	TAGAACAAGAAAATGAAGCTGTGGAGAACCTCCAGAGCACTTTGATTCAACTACAAAA	600
Oy	541	TAGAACAAGAAAATGAAGCTGTGGAGAACCTCCAGAGCACTTTGATTCAACTACAAAA	600
Db	541	TAGAACAAGAAAATGAAGCTGTGGAGAACCTCCAGAGCACTTTGATTCAACTACAAAA	600
Oy	601	CCCTCAAGAGGCAAGGAGACATGCAGAGTCTGAATGGAACAACCAAGTCTGTGACCAAGC	660
Oy	601	CCCTCAAGAGGCAAGGAGACATGCAGAGTCTGAATGGAACAACCAAGTCTGTGACCAAGC	660
Db	601	CCCTCAAGAGGCAAGGAGACATGCAGAGTCTGAATGGAACAACCAAGTCTGTGACCAAGC	660
Oy	661	AGAAGATGCAAGCAGCTGGAACAGATGCTCAAGACCCCTGGACCAGATGGCGAAGACATTG	720
Oy	661	AGAAGATGCAAGCAGCTGGAACAGATGCTCAAGACCCCTGGACCAGATGGCGAAGACATTG	720
Db	661	AGAAGATGCAAGCAGCTGGAACAGATGCTCAAGACCCCTGGACCAGATGGCGAAGACATTG	720
Oy	721	TGAGTGAAGCTGGCGGGGCTTGTGACGAATGAGATAGCTGAGAAAGACACTGTGATG	780
Oy	721	TGAGTGAAGCTGGCGGGGCTTGTGACGAATGAGATAGCTGAGAAAGACACTGTGATG	780
Db	721	TGAGTGAAGCTGGCGGGGCTTGTGACGAATGAGATAGCTGAGAAAGACACTGTGATG	780
Oy	781	AAGAGCTGGGCTGACTGGAAAGAGCGGCCAGAGATCGCTGATCTGGAGGGCCCTCCCAACA	840
Oy	781	AAGAGCTGGGCTGACTGGAAAGAGCGGCCAGAGATCGCTGATCTGGAGGGCCCTCCCAACA	840
Db	781	AAGAGCTGGGCTGACTGGAAAGAGCGGCCAGAGATCGCTGATCTGGAGGGCCCTCCCAACA	840
Oy	841	TCTGCTGTGACCGCTCTGAAAAAATCGATACTTCAATACCAAACTCTCAAGACCC	900
Oy	841	TCTGCTGTGACCGCTCTGAAAAAATCGATACTTCAATACCAAACTCTCAAGACCC	900
Db	841	TCTGCTGTGACCGCTCTGAAAAAATCGATACTTCAATACCAAACTCTCAAGACCC	900
Oy	901	GCCAAACAAATTAAGAAACTGAGAGAGCTCAGCAGAAATGTCTACAAGGGCGACCTTA	960
Oy	901	GCCAAACAAATTAAGAAACTGAGAGAGCTCAGCAGAAATGTCTACAAGGGCGACCTTA	960
Db	901	GCCAAACAAATTAAGAAACTGAGAGAGCTCAGCAGAAATGTCTACAAGGGCGACCTTA	960
Oy	961	TCTGTGACAGCACGGCCCATGCTGTGAGAGAGATCGTGGAGCTGTTCAGAAAATTAAATGA	1020
Oy	961	TCTGTGACAGCACGGCCCATGCTGTGAGAGAGATCGTGGAGCTGTTCAGAAAATTAAATGA	1020
Db	961	TCTGTGACAGCACGGCCCATGCTGTGAGAGAGATCGTGGAGCTGTTCAGAAAATTAAATGA	1020

Qy	1021	AGAGTGGCTTCTGTGTGAGCGGACGCTTCGATGCGCCATGACACCGGACCGGCTTAC	1080
Db	1021	AGATGCTCTTCGTGTGTGAGCGGACCGCTTCGATGCGCCATGACACCGGACCGGCTTAC	1080
Qy	1081	TCATCAGACTGGTGTCCATTTACGACGAAGAAGTCAAGTGTGTGTCAAAATTTCTGAT	1140
Db	1081	TCATCAGACTGGTGTCCATTTACGACGAAGAAGTCAAGTGTGTGTCAAAATTTCTGAT	1140
Qy	1141	TGAATATTCAGCTTTAAATTTAAAGTGTGCATTTAAACATCTTCGGATCTCTCC	1200
Db	1141	TGAATATTCAGCTTTAAATTTAAAGTGTGCATTTAAACATCTTCGGATCTCTCC	1200
Qy	1201	TCAGAGGCTCCGGAATTTTAACTTCTGGGACGCAACCAAGCTGTGAATCTGAGG	1260
Db	1201	TCAGAGGCTCCGGAATTTTAACTTCTGGGACGCAACCAAGCTGTGAATCTGAGG	1260
Qy	1261	AGCTTACCAACGGCAGCCTGTCTGCAGAGTTCAAGCAGCTGACCTTAGAGGACAGAT	1320
Db	1261	AGCTTACCAACGGCAGCCTGTCTGCAGAGTTCAAGCAGCTGACCTTAGAGGACAGAT	1320
Qy	1321	GTGGAAATGGAGCGCCGTGCCAATTTGTATGCCCTCTTGATCGAGCTGGAGCGCAC	1380
Db	1321	GTGGAAATGGAGCGCCGTGCCAATTTGTATGCCCTCTTGATCGAGCTGGAGCGCAC	1380
Qy	1381	TGATACACCTTCGAGACCTGAGGTGTACCAACCAAGGCTCAAGATTGACCTAGACCACT	1440
Db	1381	TGATACACCTTCGAGACCTGAGGTGTACCAACCAAGGCTCAAGATTGACCTAGACCACT	1440
Qy	1441	CCTTGCACGTTGTGTGATCTCCACACATCTGTAGATGCGCAATGCTTGGCATCAATCC	1500
Db	1441	CCTTGCACGTTGTGTGATCTCCACACATCTGTAGATGCGCAATGCTTGGCATCAATCC	1500
Qy	1501	TGTGTATTAACATGCTGTGACCAATTAACCCCAAGAACCTGAACTTCTTCACTAAAGCCGCA	1560
Db	1501	TGTGTATTAACATGCTGTGACCAATTAACCCCAAGAACCTGAACTTCTTCACTAAAGCCGCA	1560
Qy	1561	TTGGAACCTGGACCAAGTGGCGAGGTGCTCAGCTGCGCATTTCTGTCCACACCAAGC	1620
Db	1561	TTGGAACCTGGACCAAGTGGCGAGGTGCTCAGCTGCGCATTTCTGTCCACACCAAGC	1620
Qy	1621	GAGGGCTAGCATGAGCAGCTGTGACCAAGCGTGGCTGAGAAAGTCTTGAAGGCTGTGTGA	1680
Db	1621	GAGGGCTAGCATGAGCAGCTGTGACCAAGCGTGGCTGAGAAAGTCTTGAAGGCTGTGTGA	1680
Qy	1681	ACTACTCAGGGTGTTCAGATCAGATGGCTAAATTTGTGCAAGAAGAAACATGCTGCGCAAG	1740
Db	1681	ACTACTCAGGGTGTTCAGATCAGATGGCTAAATTTGTGCAAGAAGAAACATGCTGCGCAAG	1740
Qy	1741	GCTTCTCTTCTGGGTCTGGCTAACAAATATCATCACTTGTGAAAGATTAATCTTGG	1800
Db	1741	GCTTCTCTTCTGGGTCTGGCTAACAAATATCATCACTTGTGAAAGATTAATCTTGG	1800
Qy	1801	CCCTTTGGAATGAAGGATACATCATGTGGTTTCATCAGACAGAGCGGAGCGGCCATCC	1860
Db	1801	CCCTTTGGAATGAAGGATACATCATGTGGTTTCATCAGACAGAGCGGAGCGGCCATCC	1860
Qy	1861	TAAAGCAACAAGCCCCCGGGACCTTCTACTGGCTTCACCGAGACACCAAAAGAGAG	1920
Db	1861	TAAAGCAACAAGCCCCCGGGACCTTCTACTGGCTTCACCGAGACACCAAAAGAGAG	1920
Qy	1921	GGGTCACTTTCATCTGGGTGGAAGAAAGACATCAGTGGCAAGACCCAGATCCAGTCTGAG	1980
Db	1921	GGGTCACTTTCATCTGGGTGGAAGAAAGACATCAGTGGCAAGACCCAGATCCAGTCTGAG	1980
Qy	1981	AGCCATACACCAAGCAGCAGCTGGAACAACATGTCTATTGTGAAATCATCATGGCTATA	2040
Db	1981	AGCCATACACCAAGCAGCAGCTGGAACAACATGTCTATTGTGAAATCATCATGGCTATA	2040
Qy	2041	AGATCATGATGCGACCAACATCTGTGTCTCCACTTGTCTACCTTACCCGACATTC	2100
Db	2041	AGATCATGATGCGACCAACATCTGTGTCTCCACTTGTGTCTACCTTACCCGACATTC	2100

OY 2101 CCAAGAGGAGGACATTTGAAAGTACTGTAGGCCCGAGAGCCAGACCCGGAAGCCG 2160
|||||
Db 2101 CCAAGAGGAGGACATTTGAAAGTACTGTAGGCCCGAGAGCCAGACCCGGAAGCCG 2160
OY 2161 ACCCAGGTGCTGCTGCCCCGTACCTGGAAGCCAGTTCACTCTGTGACCCCAACACCT 2220
|||||
Db 2161 ACCCAGGTGCTGCTGCCCCGTACCTGGAAGCCAGTTCACTCTGTGACCCCAACACCT 2220
OY 2221 GCAGCAATACCATTTGACCTGCGCATGTCCCGCCGACCTTTAGATTCATTTGATGATG 2280
|||||
Db 2221 GCAGCAATACCATTTGACCTGCGCATGTCCCGCCGACCTTTAGATTCATTTGATGATG 2280
OY 2281 GAAATTAACGGTGAAGTGTGAGCCCTCAGCAGAGAGGAGGAGTTGATGCTGCTGCTG 2340
|||||
Db 2281 GAAATTAACGGTGAAGTGTGAGCCCTCAGCAGAGAGGAGGAGTTGATGCTGCTGCTGCTG 2340
OY 2341 ACATGATGCTGACCTGCGAGTGTGCTACCTCCCGCATGTGAGAGGCTGAAACAGAGCT 2400
|||||
Db 2341 ACATGATGCTGACCTGCGAGTGTGCTACCTCCCGCATGTGAGAGGCTGAAACAGAGCT 2400
OY 2401 GCAGAGACGCTGACCTGAGACACCTGCGCGCTGCTCACCCTTAAGCAGCCGAAACCCATA 2460
|||||
Db 2401 GCAGAGACGCTGACCTGAGACACCTGCGCGCTGCTCACCCTTAAGCAGCCGAAACCCATA 2460
OY 2461 TCGTCTGAACCTCTTACCTTTGTGCTTCCAGATTTTTTTTTTAATTTCTCTACTTCTGCT 2520
|||||
Db 2461 TCGTCTGAACCTCTTACCTTTGTGCTTCCAGATTTTTTTTTTAATTTCTCTACTTCTGCT 2520
OY 2521 ATCTTTGGCAATCTGGGCACTTTTAAAGAGAGAAATGAGTGTGAGTGTGATTAAC 2580
|||||
Db 2521 ATCTTTGGCAATCTGGGCACTTTTAAAGAGAGAAATGAGTGTGAGTGTGATTAAC 2580
OY 2581 TGTATGTAAAGAGAGAGACCTCTGAGTCTGGGATGGGGCTGAGAGCAGAAAGGAGGC 2640
|||||
Db 2581 TGTATGTAAAGAGAGAGACCTCTGAGTCTGGGATGGGGCTGAGAGCAGAAAGGAGGC 2640
OY 2641 AAGGAGGAGACACCTCTGCTGCTGCGCGCTGCTCTTTTTCAGAGAGCTGGGGGTTGG 2700
|||||
Db 2641 AAGGAGGAGACACCTCTGCTGCTGCGCGCTGCTCTTTTTCAGAGAGCTGGGGGTTGG 2700
OY 2701 TTGTTAGACAACTGCTCTGCTGCTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
|||||
Db 2701 TTGTTAGACAACTGCTCTGCTGCTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
OY 2761 ACCCATCTGAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
|||||
Db 2761 ACCCATCTGAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
OY 2821 TAGGACTAAGCCAGAGAGGTTCTCTTTAAATTTAAAAA 2869
|||||
Db 2821 TAGGACTAAGCCAGAGAGGTTCTCTTTAAATTTAAAAA 2869

RESULT 2
LOCUS 187790 2869 bp DNA PAT 10-AUG-1998
DEFINITION Sequence 11 from patent US 5716622.
ACCESSION 187790
VERSION 187790.1 GI:3407730
KEYWORDS
ORGANISM Unknown.
SOURCE Unknown.
REFERENCE 1 (bases 1 to 2869)
AUTHORS Darnell,J.E. Jr., Wen,Z., Horvath,C.M. and Zhong,Z.
TITLE Functionally active regions of signal transducer and activators of transcription
Patent: US 5716622-A 11 10-FEB-1998;
JOURNAL Location/Qualifiers
FEATURES
SOURCE 1..2869
BASE COUNT 755 a 743 c 772 g 599 t
ORIGIN

Query Match 100.0%: Score 2869; DB 6; Length 2869;
Best Local Similarity 100.0%: Pred No. 0;
Matches 2869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCCCGACACAGCCAGCCGCGCTCAGCCGCGGAGACAGTGCAGACCCCTGACT 60
|||||
Db 1 GCCCGACACAGCCAGCCGCGCTCAGCCGCGGAGACAGTGCAGACCCCTGACT 60
OY 61 GCACAGAGATGGCTCATGTGGAACCACTGACAGAGCTGACACACCTACTGAAAGAC 120
|||||
Db 61 GCACAGAGATGGCTCATGTGGAACCACTGACAGAGCTGACACACCTACTGAAAGAC 120
OY 121 TGACACAGCTGTACAGAGACAGCTTCCCATGAGCTGCGGCAATTTCTTGACCTTTGA 180
|||||
Db 121 TGACACAGCTGTACAGAGACAGCTTCCCATGAGCTGCGGCAATTTCTTGACCTTTGA 180
OY 181 TTGAGAGTCAAGACTGGGCAATATGACCCAGCAAAAGATGACATGCCACGTTGGTTTC 240
|||||
Db 181 TTGAGAGTCAAGACTGGGCAATATGACCCAGCAAAAGATGACATGCCACGTTGGTTTC 240
OY 241 ATATATCTTTGGGGAATGATGACCAATATATGCGATTTCTGCAAGAGATCCATATGTC 300
|||||
Db 241 ATATATCTTTGGGGAATGATGACCAATATATGCGATTTCTGCAAGAGATCCATATGTC 300
OY 301 TCTATCAGCACAACCTTTCGAAGAATCAAGCAGTTTCTGCAGAGCAGGATCTTGAGAAC 360
|||||
Db 301 TCTATCAGCACAACCTTTCGAAGAATCAAGCAGTTTCTGCAGAGCAGGATCTTGAGAAC 360
OY 361 CAATGCAATTTGCCCGGATCTGCGCCGATGCTGTGGGAAGTCTGCGCTCTCCAGA 420
|||||
Db 361 CAATGCAATTTGCCCGGATCTGCGCCGATGCTGTGGGAAGTCTGCGCTCTCCAGA 420
OY 421 CGCAGCGACAGGAGCCAGCCAGAGGGGGCCAGGCCAACCCCAACCGCCGCTGATGA 480
|||||
Db 421 CGCAGCGACAGGAGCCAGCCAGAGGGGGCCAGGCCAACCCCAACCGCCGCTGATGA 480
OY 481 CAGAGAGCAGCAGATGTTGGAGCAGCATCTTGAGATGTCGCGAAGCAGTGCAGAGATC 540
|||||
Db 481 CAGAGAGCAGCAGATGTTGGAGCAGCATCTTGAGATGTCGCGAAGCAGTGCAGAGATC 540
OY 541 TAAACAGAAATGAAGCTGTGGAGAACCTCCAGAGCAGCTTGTGATTTAACTACAAAA 600
|||||
Db 541 TAAACAGAAATGAAGCTGTGGAGAACCTCCAGAGCAGCTTGTGATTTAACTACAAAA 600
OY 601 CCCTCAAGAGCCAGAGAGATGACAGATCTGAATGGAACCAACCAAGTGTGACCCAGAC 660
|||||
Db 601 CCCTCAAGAGCCAGAGAGATGACAGATCTGAATGGAACCAACCAAGTGTGACCCAGAC 660
OY 661 AGAAGATGACAGCAGCTGGAACAGATGCTCACAGCCCTGAGCCAGATGCGGAGAAGCATTTG 720
|||||
Db 661 AGAAGATGACAGCAGCTGGAACAGATGCTCACAGCCCTGAGCCAGATGCGGAGAAGCATTTG 720
OY 721 TGAGTAGCTGGGGGGGCTTTGTCAAGCAATGAGTGTGCAAGAGACACTGACTGATG 780
|||||
Db 721 TGAGTAGCTGGGGGGGCTTTGTCAAGCAATGAGTGTGCAAGAGACACTGACTGATG 780
OY 781 AAGAGCTGGCTGACTGGAAGAGCGGCGCAGAGATGCGTGCATGCGAGGCGCTCCCAACA 840
|||||
Db 781 AAGAGCTGGCTGACTGGAAGAGCGGCGCAGAGATGCGTGCATGCGAGGCGCTCCCAACA 840
OY 841 TCTGCTTGACCGCTTGAAAACTGATTAATTTAGTACAGATCTCAACTTCACAGACC 900
|||||
Db 841 TCTGCTTGACCGCTTGAAAACTGATTAATTTAGTACAGATCTCAACTTCACAGACC 900
OY 901 GCCAACAATTTAGAACTGAGAGAGCTGACAGAGAAATGTCTTCAAGAGGCGACCCCTA 960
|||||
Db 901 GCCAACAATTTAGAACTGAGAGAGCTGACAGAGAAATGTCTTCAAGAGGCGACCCCTA 960
OY 961 TCGTGCAGCAGCGGCGCATGCTGAGAGAGAGATGCTGAGAGCTGTTCAAGAACTTAATGA 1020
|||||
Db 961 TCGTGCAGCAGCGGCGCATGCTGAGAGAGAGATGCTGAGAGCTGTTCAAGAACTTAATGA 1020

1021 AGAGTCCCTTCTGTTGGAGCGGACCCCTGCATGCCCATGCACCCGAGCCGCTTAC 1080
1021 AGAGTCCCTTCTGTTGGAGCGGACCCCTGCATGCCCATGCACCCGAGCCGCTTAC 1080
1081 TCATCAAGACTGGTGTCCACTTTACCGCAAGCTCAGCTGGTGTCTAAATTTCTGAGT 1140
1081 TCATCAAGACTGGTGTCCACTTTACCGCAAGCTCAGCTGGTGTCTAAATTTCTGAGT 1140
1141 TGAATTAATCAAGTAAATTAAGTGTGATTAAGACTCCTGGGATGTTGCTGCC 1200
1141 TGAATTAATCAAGTAAATTAAGTGTGATTAAGACTCCTGGGATGTTGCTGCC 1200
1201 TCAGAGGCTTCGGAATTTAACTTTGGGCGACGAACAAAAAGTATGAACATGGAGG 1260
1201 TCAGAGGCTTCGGAATTTAACTTTGGGCGACGAACAAAAAGTATGAACATGGAGG 1260
1261 AGCTTAACAAGCGAGCGCTCTGCGAGATTCAAGACCTGACCTTAGGAGCAGAGAT 1320
1261 AGCTTAACAAGCGAGCGCTCTGCGAGATTCAAGACCTGACCTTAGGAGCAGAGAT 1320
1321 GTGGGAATGGAGCGCTGCCAATTTGATGCTCTTGTATGTCGATGAGAGCTGCACC 1380
1321 GTGGGAATGGAGCGCTGCCAATTTGATGCTCTTGTATGTCGATGAGAGCTGCACC 1380
1381 TGATCACTTCGAGACTGAGGTGTACCAAGGCTCAAGATTGACCTAGAGCCACT 1440
1381 TGATCACTTCGAGACTGAGGTGTACCAAGGCTCAAGATTGACCTAGAGCCACT 1440
1441 CCTTGCAGTTGTGTGATCTCCACATCTGTGATGATGATGATGATGATGATGATGAT 1500
1441 CCTTGCAGTTGTGTGATCTCCACATCTGTGATGATGATGATGATGATGATGATGAT 1500
1501 TGTGATTAATCAAGTAAATTAAGTGTGATTAAGACTCCTGGGATGTTGCTGCC 1560
1501 TGTGATTAATCAAGTAAATTAAGTGTGATTAAGACTCCTGGGATGTTGCTGCC 1560
1561 TTGGAACCTGGAGCAAGTGGCGGAGCTGCTCAGCTGGAGTTCTGTCACACCAAGC 1620
1561 TTGGAACCTGGAGCAAGTGGCGGAGCTGCTCAGCTGGAGTTCTGTCACACCAAGC 1620
1621 GAAGGCTGAGACTGAGAGCTGACCAAGCTGCTGAGAAAGCTCTTAGGGCTGTGTGA 1680
1621 GAAGGCTGAGACTGAGAGCTGACCAAGCTGCTGAGAAAGCTCTTAGGGCTGTGTGA 1680
1681 ACTACTGAGGCTGAGAGCTGACCAAGCTGCTGAGAAAGCTCTTAGGGCTGTGTGA 1740
1681 ACTACTGAGGCTGAGAGCTGACCAAGCTGCTGAGAAAGCTCTTAGGGCTGTGTGA 1740
1741 GCTTCTCTTCTGGGCTGTGCTAGACAAATATCATCGACTTGTGAAAAGTATATCTTGG 1800
1741 GCTTCTCTTCTGGGCTGTGCTAGACAAATATCATCGACTTGTGAAAAGTATATCTTGG 1800
1801 CCTTTGGAATGAAGGTTACATATGCTTTCATACAGCAAGGCGGAGCGGCAATCC 1860
1801 CCTTTGGAATGAAGGTTACATATGCTTTCATACAGCAAGGCGGAGCGGCAATCC 1860
1861 TAAGCAACAAAGCCCGGCGACCTTCTTACTGCGCTTACGCGAGAGCAAGAAAGAG 1920
1861 TAAGCAACAAAGCCCGGCGACCTTCTTACTGCGCTTACGCGAGAGCAAGAAAGAG 1920
1921 GGGTCACTTCTGAGTGGTGAAGAGACATCAAGTGGCAGAGACCCAGATCCTGTAG 1980
1921 GGGTCACTTCTGAGTGGTGAAGAGACATCAAGTGGCAGAGACCCAGATCCTGTAG 1980
1981 AGCCATACAAAGCAAGCAAGCTGACATGTCATTTGCTGAATCATATGAGCTATA 2040
1981 AGCCATACAAAGCAAGCAAGCTGACATGTCATTTGCTGAATCATATGAGCTATA 2040
2041 AGATCATGATGAGCAACATCTGCTGCTCACTTGTCTACCTTACCCGAGATTC 2100
2041 AGATCATGATGAGCAACATCTGCTGCTCACTTGTCTACCTTACCCGAGATTC 2100

2101 CCAAGAGAGCAATTTGAAAGTACTGTAGGCCGAGAGCCAGAGCACCACCCGAGCCG 2160
2101 CCAAGAGAGCAATTTGAAAGTACTGTAGGCCGAGAGCACCACCCGAGCCGAGCCG 2160
2161 ACCGAGTACGCTGCGCCGCTACCTGAAAGCAAGTTCATCTGTGTGACACCAAGCACT 2220
2161 ACCGAGTACGCTGCGCCGCTACCTGAAAGCAAGTTCATCTGTGTGACACCAAGCACT 2220
2221 GCACCAATACCATTTGACCTGCGGATGTCGCCCGCAGCTTATGATTCATGATGAGTTTG 2280
2221 GCACCAATACCATTTGACCTGCGGATGTCGCCCGCAGCTTATGATTCATGATGAGTTTG 2280
2281 GAAATACGCTGAAAGTGTCTGAGCCCTCAGCAGAGAGGAGCTTTGATCGCTACGTTTG 2340
2281 GAAATACGCTGAAAGTGTCTGAGCCCTCAGCAGAGAGGAGCTTTGATCGCTACGTTTG 2340
2341 ACATGATCTGACCTCGAGTGTCTACCTGCTCCCATGTAGAGAGCTGAAACAGAACT 2400
2341 ACATGATCTGACCTCGAGTGTCTACCTGCTCCCATGTAGAGAGCTGAAACAGAACT 2400
2401 GCAGAGACGTGACTTGAAGACACTGCGCGCTGCTCCACCCCTAAGCAGCCGAACTTATA 2460
2401 GCAGAGACGTGACTTGAAGACACTGCGCGCTGCTCCACCCCTAAGCAGCCGAACTTATA 2460
2461 TCGTCTGAAACTCCTTACTTGTGCTGCTCAGATTTTATTTTATTTTATTTTCTTCTGCT 2520
2461 TCGTCTGAAACTCCTTACTTGTGCTGCTCAGATTTTATTTTATTTTATTTTCTTCTGCT 2520
2521 ATCTTTGGGCAATCTGGGCACTTTTAAAGAGAAATGATGATGATGATGATGATGATGAT 2580
2521 ATCTTTGGGCAATCTGGGCACTTTTAAAGAGAAATGATGATGATGATGATGATGATGAT 2580
2581 TGTATTAATTAAGAGAGAGACCTGAGCTGCTGGGAGTGGGCTGAGAGCAAGAGGAGG 2640
2581 TGTATTAATTAAGAGAGAGACCTGAGCTGCTGGGAGTGGGCTGAGAGCAAGAGGAGG 2640
2641 AAAGGCAACACCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
2641 AAAGGCAACACCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
2701 TTGTTAGACAAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
2701 TTGTTAGACAAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
2761 ACCCATTTCTGGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
2761 ACCCATTTCTGGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
2821 TAGGACTAAGCCCAAGAGGTTCTCTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2869
2821 TAGGACTAAGCCCAAGAGGTTCTCTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2869

RESULT 3
LOCUS AR121061 2869 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 82 from patent US 6159694.
ACCESSION AR121061
VERSION AR121061.1 GI:14104637
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2869)
AUTHORS Karas,J.G.
TITLE Antisense modulation of stat3 expression
JOURNAL Patent: US 6159694-A 82 12-DEC-2000;
FEATURES
source
1..2869
/organism="unknown"
BASE COUNT 755 a 743 c 772 g 599 t
ORIGIN

Query Match	99.8%	Score 2864.2	DB 6	Length 2869	
Best Local Similarity	99.9%	Pred. No. 0			
Matches 2866	Conservative	0	Mismatches	3	Indels
					Gaps
OY	1	GCCTGACAGCAGCAGCGCCGCGCCATGTCGGGCTTCAGCCCGAGACAGTGCAGACCTCTGACT	60		
Db	1	GCCTGACAGCAGCAGCGCCGCGCCATGTCGGGCTTCAGCCCGAGACAGTGCAGACCTCTGACT	60		
OY	61	GCAGCAGATGGCTCAGTGGAAACCAAGCTGCACACCTGGACACACGCTAACCTGGAAGCAGC	120		
Db	61	GCAGCAGATGGCTCAGTGGAAACCAAGCTGCACACCTGGACACACGCTAACCTGGAAGCAGC	120		
OY	121	TGCACGAGTGTACAGCGCACGCTTCCCATGAGACTGGCGCAGTGTCCGCGACCTTGA	180		
Db	121	TGCACGAGTGTACAGCGCACGCTTCCCATGAGACTGGCGCAGTGTCCGCGACCTTGA	180		
OY	181	TTGAGAGTCAACACTGGGCTATGCAGCCAGCAAAAGATCACATGCGACGTTGGTGTTC	240		
Db	181	TTGAGAGTCAACACTGGGCTATGCAGCCAGCAAAAGATCACATGCGACGTTGGTGTTC	240		
OY	241	ATAATCTCTGGGTGAATTTGACACAGCAATTAAGCGATTCTCTGCNAGAGTCCAAATGTCC	300		
Db	241	ATAATCTCTGGGTGAATTTGACACAGCAATTAAGCGATTCTCTGCNAGAGTCCAAATGTCC	300		
OY	301	TCTATCAGCACAACTTGCAGANTCAGCAGTTCTGCAGACAGTATCTTTGAGAAGC	360		
Db	301	TCTATCAGCACAACTTGCAGANTCAGCAGTTCTGCAGACAGTATCTTTGAGAAGC	360		
OY	361	CAATGAAATTTGCCGAGTCGTGGCCCGATGCTGTGGGAAAGTTCCTCCTCTCCAGA	420		
Db	361	CAATGAAATTTGCCGAGTCGTGGCCCGATGCTGTGGGAAAGTTCCTCCTCTCCAGA	420		
OY	421	CGGACCCACAGCGACGCCACGCAAGGGGGCCAGCCAAACCAACCAAGCCCGCTAATGA	480		
Db	421	CGGACCCACAGCGACGCCACGCAAGGGGGCCAGCCAAACCAACCAAGCCCGCTAATGA	480		
OY	481	CAGAGAAAGAGCAGATGTTGGAGCACATCTTTCAGAGATGCCGGAAGCAGTGCAGATTC	540		
Db	481	CAGAGAAAGAGCAGATGTTGGAGCACATCTTTCAGAGATGCCGGAAGCAGTGCAGATTC	540		
OY	541	TAGAACAGAAATGAAGGTGTGGAGAACCTCCAGGACGACTTTGATTTCATACACAAA	600		
Db	541	TAGAACAGAAATGAAGGTGTGGAGAACCTCCAGGACGACTTTGATTTCATACACAAA	600		
OY	601	CCCTCAGAGGCCAAGGAGACATGACAGATCTGGAATGGAACCAACCGTGTGCACGAC	660		
Db	601	CCCTCAGAGGCCAAGGAGACATGACAGATCTGGAATGGAACCAACCGTGTGCACGAC	660		
OY	661	AGAAAGTGCAGCAGCTGGAACAGATCTTCACAGCCCTTGACACGATGCAGAAAGCATTC	720		
Db	661	AGAAAGTGCAGCAGCTGGAACAGATCTTCACAGCCCTTGACACGATGCAGAAAGCATTC	720		
OY	721	TGAGTACAGCTGGCGGGGCTTGTCAAGCAATGAGATGATGTCGGAAGAACATGACTGATG	780		
Db	721	TGAGTACAGCTGGCGGGGCTTGTCAAGCAATGAGATGATGTCGGAAGAACATGACTGATG	780		
OY	781	AAGAGTGGCTGCATGGAAGAGCGGCGCACAGATCCGCTGCATCGGAGGCGCTCCCAACA	840		
Db	781	AAGAGTGGCTGCATGGAAGAGCGGCGCACAGATCCGCTGCATCGGAGGCGCTCCCAACA	840		
OY	841	TCTGCTGAGCAGCTGTGGAATACTGATTAATCTTATAGCAGATTTCAACTTCAGACCC	900		
Db	841	TCTGCTGAGCAGCTGTGGAATACTGATTAATCTTATAGCAGATTTCAACTTCAGACCC	900		
OY	901	GCCAAACAAATTAGAACTGGAGAGCTGCAGAGAAAGTGTCTTCAAGGGGAGCCCTA	960		
Db	901	GCCAAACAAATTAGAACTGGAGAGCTGCAGAGAAAGTGTCTTCAAGGGGAGCCCTA	960		
OY	961	TGTGTGAGCAGCGGCCCATGCTGTGGAGAGAGATGCTGAGCTGTTTCAGAACTTAATGA	1020		
Db	961	TGTGTGAGCAGCGGCCCATGCTGTGGAGAGAGATGCTGAGCTGTTTCAGAACTTAATGA	1020		

QY	1021	AGATGCTCTTCGTGCTGTGAGCGGACGCCACTCGATGCGCATGACACCGGGACCGGCCCTTAG	1080
Db	1021	AGATGCTCTTCGTGCTGTGAGCGGACCGGACGCCCTCGATGCGCATGACACCGGGACCGGCCCTTAG	1080
QY	1081	TCATCAAGACCTGGGTGCTCAGTTTACCACCAAAAGTCAGGTGCTGTGTCAAATTTCTAGT	1140
Db	1081	TCATCAAGACCTGGGTGCTCAGTTTACCACCAAAAGTCAGGTGCTGTGTCAAATTTCTAGT	1140
QY	1141	TGAATATACAGCTTTAAATTTAAAGTGTGCATTTGATTAAGACTCTGTGGGATGTTGCTGCC	1200
Db	1141	TGAATATACAGCTTTAAATTTAAAGTGTGCATTTGATTAAGACTCTGTGGGATGTTGCTGCC	1200
QY	1201	TCAGAGGCTCTCGGAATTTAAACATTTCTGGGACGCAACCAAAAGTGAATGAACATGAGAG	1260
Db	1201	TCAGAGGCTCTCGGAATTTAAACATTTCTGGGACGCAACCAAAAGTGAATGAACATGAGAG	1260
QY	1261	AGCTC...CAACGGGAGGCTGTCTGCAAGTTCGAAGCACCCTGACCCCTTAAGGAGCGAGAGAT	1320
Db	1261	AGCTC...CAACGGGAGGCTGTCTGCAAGTTCGAAGCACCCTGACCCCTTAAGGAGCGAGAGAT	1320
QY	1321	GTTGGGAATGGAAGCGCCGGGCCAATTTGGTATGCGCCCTTGATTCGAGCTCAAGAGCTGAC	1380
Db	1321	GTTGGGAATGGAAGCGCCGGGCCAATTTGGTATGCGCCCTTGATTCGAGCTCAAGAGCTGAC	1380
QY	1381	TGATACACCTTCGAGACTGAGAGTGTACACCAAGGCCCTCAAGATTGACCTAGAGACCCACT	1440
Db	1381	TGATACACCTTCGAGACTGAGAGTGTACACCAAGGCCCTCAAGATTGACCTAGAGACCCACT	1440
QY	1441	CCTTGCCAGTTGTGTGTATCTCCACACATCTGTAGATGCGCAATGCTTTGGGATCAATCC	1500
Db	1441	CCTTGCCAGTTGTGTGTATCTCCACACATCTGTAGATGCGCAATGCTTTGGGATCAATCC	1500
QY	1501	TGTGCTTAACATCTGACCAATTAACCCCAAGAACGTGAACCTTCTCACTAAGCGGCCAA	1560
Db	1501	TGTGCTTAACATCTGACCAATTAACCCCAAGAACGTGAACCTTCTCACTAAGCGGCCAA	1560
QY	1561	TTGGAACCTGGGACCAAGTGGCGAGTGTCAGCTGGCAGTTCTCTGCACACCAACGC	1620
Db	1561	TTGGAACCTGGGACCAAGTGGCGAGTGTCAGCTGGCAGTTCTCTGCACACCAACGC	1620
QY	1621	GAGGGCTGAGCATGAGACACGCTGCAACGCTGGCTAGAAAGCCTCAGGGCCTGTGTGA	1680
Db	1621	GAGGGCTGAGCATGAGACACGCTGCAACGCTGGCTAGAAAGCCTCAGGGCCTGTGTGA	1680
QY	1681	ACTACTCAGGGTGTCAAGATCACATGAGGCTAAATTTGCAAAAGAAACATGCTGGCAAG	1740
Db	1681	ACTACTCAGGGTGTCAAGATCACATGAGGCTAAATTTGCAAAAGAAACATGCTGGCAAG	1740
QY	1741	GCTTCTCCTTCTGGGCTGTGGCTAACAATTCATCGACCTTGTGAAGAAAGTATATCTTGG	1800
Db	1741	GCTTCTCCTTCTGGGCTGTGGCTAACAATTCATCGACCTTGTGAAGAAAGTATATCTTGG	1800
QY	1801	CCCTTTGGAATGAGAGGTATCATCTGAGGTTTCATCAGCAAGGAGCGGAGCGGCATCC	1860
Db	1801	CCCTTTGGAATGAGAGGTATCATCTGAGGTTTCATCAGCAAGGAGCGGAGCGGCATCC	1860
QY	1861	TAAAGCACAAGCCCCCGGGCACTCTTCACTGGCCTTCACCGAGACACCAAGAAAGAG	1920
Db	1861	TAAAGCACAAGCCCCCGGGCACTCTTCACTGGCCTTCACCGAGACACCAAGAAAGAG	1920
QY	1921	GGGTACACTTTCATCTGGGTGGAAAAAGACATCAGTGGCAAGACCCGAGTCCAGTGTGAG	1980
Db	1921	GGGTACACTTTCATCTGGGTGGAAAAAGACATCAGTGGCAAGACCCGAGTCCAGTGTGAG	1980
QY	1981	AGCCATACCAAGCAGCAGCTGAACAACATGTCAATTTGCTGAATCATATGSGCTATA	2040
Db	1981	AGCCATACCAAGCAGCAGCTGAACAACATGTCAATTTGCTGAATCATATGSGCTATA	2040
QY	2041	AGATATGATGGAGCAACAATCCTGGTGTCTCACTGTCTACCTCTACCCCGCATTC	2100
Db	2041	AGATATGATGGAGCAACAATCCTGGTGTCTCACTGTCTACCTCTACCCCGCATTC	2100
QY	2101	CCAAGGAGAGCAATTTTGAAAGTACTGTAGGCCCGGAGACCAGGAGCACCCCGAAGCGG	2160

Db	2101	CCAAAGGAGGAGCATTTTGGAAAGTACTGTAGGCCCGGAGAGCCAGAGGACACCCCGAAGCCG	2160
QY	2161	ACCCAGGTAGTGTGCGCCCGTACTGTGAAGACCAAGTGTATCTGTGTGACACCAAGCACT	2220
Db	2161	ACCCAGGTAGTGTGCGCCCGTACTGTGAAGACCAAGTGTATCTGTGTGACACCAAGCACT	2220
QY	2221	GTACCAATACCATTGACCTGTCCGATGTGTCCCGCCGACCTTGTAGATTCATGTATGCAAGTTTG	2280
Db	2221	GCACCAATACCATTGACCTGTCCGATGTGTCCCGCCGACCTTGTAGATTCATGTATGCAAGTTTG	2280
QY	2281	GAATTAACGGGTGAAGGTGTGAGACCCCTCAGACAGAGAGGCGAGTTTGTAGTCCGTCAAGTTTG	2340
Db	2281	GAATTAACGGGTGAAGGTGTGAGACCCCTCAGACAGAGAGGCGAGTTTGTAGTCCGTCAAGTTTG	2340
QY	2341	ACATGATCTGTGACCTGTGGAGTGTGTCTACCTCCCCATGTGAGGAGCTTAAACCAGAAAGCT	2400
Db	2341	ACATGATCTGTGACCTGTGGAGTGTGTCTACCTCCCCATGTGAGGAGCTTAAACCAGAAAGCT	2400
QY	2401	GCACAGACGTGACCTGTGAGACCTGTGCCCGGTGTCCACCCCTCAACCCAGCCAGACCCCATTA	2460
Db	2401	GCACAGACGTGACCTGTGAGACCTGTGCCCGGTGTCCACCCCTCAACCCAGCCAGACCCCATTA	2460
QY	2461	TGCTGTGAAACTCCTAACTTTGTGTTCACAGATTTTTTTTTTAAATTTCTACTTCTGCT	2520
Db	2461	TGCTGTGAAACTCCTAACTTTGTGTTCACAGATTTTTTTTTTAAATTTCTACTTCTGCT	2520
QY	2521	ATCTTTGGGCATCTGTGGGCATTTTTTAAAGAGAGAAATGAGTGTGGGTGATTAAC	2580
Db	2521	ATCTTTGGGCATCTGTGGGCATTTTTTAAAGAGAGAAATGAGTGTGGGTGATTAAC	2580
QY	2581	TGTTATGTAAAGAGAGAGACACTCTGTAGTCTGTGGGTGAGAGCTGAGACAGAGAAGAGAGC	2640
Db	2581	TGTTATGTAAAGAGAGAGACACTCTGTAGTCTGTGGGTGAGAGCTGAGACAGAGAAGAGAGC	2640
QY	2641	AAAGGGGAACACCTCTGTCTCTGCGCCGTCTCTCTTTTTCAGACAGCTGGGGGGTTGG	2700
Db	2641	AAAGGGGAACACCTCTGTCTCTGCGCCGTCTCTCTCTTTTTCAGACAGCTGGGGGGTTGG	2700
QY	2701	TTGTGTAGACAAGTCCCTGTGGGTGGGCCCATGGGTACCTGTGTGCCGACGTGTGGAGCTGAT	2760
Db	2701	TTGTGTAGACAAGTCCCTGTGGGTGGGCCCATGGGTACCTGTGTGCCGACGTGTGGAGCTGAT	2760
QY	2761	ACCCCATCTTGGGAACCTCTGTGGCTGTGCACTTTCACACTTGTCTAATATTCACATAGAAAGC	2820
Db	2761	ACCCCATCTTGGGAACCTCTGTGGCTGTGCACTTTCACACTTGTCTAATATTCACATAGAAAGC	2820
QY	2821	TAGAGCTAAGCCACAGAGAGTTCTCTTTTAAATTTAAAAAAAAAAAAAAAAAAAAA	2889
Db	2821	TAGAGCTAAGCCACAGAGAGTTCTCTTTTAAATTTAAAAAAAAAAAAAAAAAAAAA	2889

RESULT	4			
LOCUS	MMU06922			
DEFINITION	MMU06922	2869 bp	mRNA	01-JUL-1994
ACCESSION	U06922			
VERSION	U06922.1	GI:458705		
KEYWORDS				
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Enayayola; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murineae; Mus			
AUTHORS	zhong,Z., Men,Z., and Darnell,J.E.			
TITLE	Stat3, a STAT family member activated by tyrosine phosphorylation in response to epidermal growth factor and interleukin-6			
JOURNAL	Science	264,	95-98	(1994)
MEDLINE	9418718			
REFERENCE	2 (bases 1 to 2869)			
AUTHORS	zhong,Z., Men,Z., and Darnell,J.			
TITLE	Stat3 and Stat5: members of the family of signal transducers and			

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
10021, USA

activators of transcription
Proc. Natl. Acad. Sci. U.S.A. 91, 4806-4810 (1994)
94255416
3 (bases 1 to 2869)
Zhong, Z.
Direct Submission
Submitted (18-FEB-1994) Z. Zhong, The Rockefeller University, The
Molecular Cell Biology Laboratory, 1230 York Avenue, New York, NY

FEATURES	Location/Qualifiers
source	1..2869
	/organism="Mus musculus"
	/db_xref="taxon:10090"
	/sex="female"
	/tissue_type="thymus"
	/clone_idb="Mouse thymus cdna library (Stratagene)"
	/dev_stage="6-8 weeks"
gene	69..2381
	/gene="Stat3"
CDS	69..2381

```

/gene="Stat3"
/function="DNA binding protein induced by EGF and IL-6"
/citation=[1]
/citation=[2]
/codon_start=1
/evidence=experimental
/product="Stat3"
/protein_id="AA19452.1"
/db_xref="GI:458706"
/translation="MAOMNLOOLDPMLDRVYLKHLQLYSDTFPPEMLRQFLPAWIESOD
YAASKESHATLVFHNHLLGIDDOYSRFLEDSNVLYOHNLRRJIKOFLOSYLEKPEE
RIVARCIEMSESLLOTAATPAAGOGGANHPATAAVYETKOOMLDHLQVRRKVOGL
KMYAVLWLDPPDPNPKTKLKSQMDMLDGNNSVYTRKOMQLEHMLATLQMRSS
SELAVGLLNAMEIVQKLTDLBELADMKRKQIACIGAPPNICLRLENNITSLAESQ
TRQDIKKEELQOKVSKGDPVQHRPMLSEERLIVELNMLSAVSEVSPOMPH
RPLIVTKGVELOFTRVLYLKFPELNTQLKIVCINDSDVAALGSRFNLTGYN
VMNDEENNGSLAEAFKHLTLRQRCGNGRANCASLIYTELHLITFEVEVYHQ
KIDITETSLPVVYISNICOMPNNAMAILWYNNMLTNPNKRVNFFETPPIGTMOVQE
SMOFSSTTKRGLSIEDLTFLAEKLRGPNVSSCOITYAKPKENNAKGSGEFNMY
NIIDLYKKYILALMNEGYIMGFISKEKEBALISTKRGPEFLRFESSSEGGVFPT
EKDISCKTQIQSVEPTTKQOLNNMSFAEILIMCKLIDATNNILYSPLYVYDIPKEE
FGKCRPSGDSHPEDPQSAAPLYLKNFKICVPTTSCNTIDLPMSPRLTDSLMOFG
GEGAEPSAGQSFSLTMDLTSECATSP"

```

Query Match	99.8%: Score 2864, 2;	DB 10; Length 2869;
Best Local Similarity	99.9%: Pred. No. 0;	
Matches 2866; Conservative	0; Mismatches	3; Indels 0; Gaps
QY 1 GCCGCGACCGACCGAGCGCGCGACGTGCGGCTCTACGCCCGAGACAGTGTGAACCCCTGACT 60		
Db 1 GCCGCGACCGACCGAGCGCGCGCGACGTGCGGCTCTACGCCCGAGAGAGTGTGAACCCCTGACT 60		
QY 61 GCAGCAGATGGCTTAGTGTGAACCAAGCTGCACGACTGTGACACACGCTACCTTAACAGAC 120		
Db 61 GCAGCAGATGGCTTAGTGTGAACCAAGCTGCACGACTGTGACACACGCTACCTTAACACAG 120		
QY 121 TGCACCACTGTACAGCGCACACGTTCCCATGTGAGAGCTGCGGCAATTCCTGGACCTTTGGA 180		
Db 121 TGCACCACTGTACAGCGCACACGTTCCCATGTGAGAGCTGCGGCAATTCCTGGACCTTTGGA 180		
QY 181 TTGAGAGTCAAGACTGGGCAATATGCAAGCCAGACAAAGATCAATGCCACGTTGGTGTTC 240		
Db 181 TTGAGAGTCAAGACTGGGCAATATGCAAGCCAGACAAAGATCAATGCCACGTTGGTGTTC 240		
QY 241 ATAATCCTTTGGGTGAATTTGACACAGCAATTTAGCCCAATTCCTGCAAGAATCCAAATGTC 300		
Db 241 ATAATCCTTTGGGTGAATTTGACACAGCAATTTAGCCCAATTCCTGCAAGAATCCAAATGTC 300		
QY 301 TCTATCAGCACACCTTGTGAAGAAATCAAGCAAGTTCCTGCGAGAGCAGTATCTTGAAGAC 360		
Db 301 TCTATCAGCACACCTTGTGAAGAAATCAAGCAAGTTCCTGCGAGAGCAGTATCTTGAAGAC 360		

```
Db 301 TCATATGACCAACCTTCGAAAGAAATCAAGCAATTCTGCGAGAGAGTAGTATCTTACGAAC 360
Oy 361 CAATGAAATTTGCCCGGATGCTGGCCCGGATGCTGTGGGAGAGTCTGGCTCTCCAGAG 420
    |||||
Db 361 CAATGAAATTTGCCCGGATGCTGGCCCGGATGCTGTGGGAGAGTCTGGCTCTCCAGAG 420
Oy 421 CGGAGCGCCAGCGGACGCGGAGAGGGGGCCAGGCCAACCAACCAAGCCGCCCTACTGA 480
    |||||
Db 421 CGGAGCGCCAGCGGACGCGGAGAGGGGGCCAGGCCAACCAACCAAGCCGCCCTACTGA 480
Oy 481 CAGAGAAAGCAGCATGTGTGGAGCAGCATCTTCAGATGTCCGAGACGAGTGCAGATC 540
    |||||
Db 481 CAGAGAAAGCAGCATGTGTGGAGCAGCATCTTCAGATGTCCGAGAGGAGATGCAGATC 540
Oy 541 TAGAAGCAAAAAATGAGGTGGTGGAGAACTCCAGACGACTTTGATTTCACTACAAA 600
    |||||
Db 541 TAGAAGCAAAAAATGAGGTGGTGGAGAACTCCAGACGACTTTGATTTCACTACAAA 600
Oy 601 CCCCAAGAGCCAGAGACATGAGATCGAGATCGAATGGAACACGAGTGTGACCAAGC 660
    |||||
Db 601 CCCCAAGAGCCAGAGACATGAGATCGAGATCGAATGGAACACGAGTGTGACCAAGC 660
Oy 661 AGAAGATGACAGCAGCTGGAAACAGATGCTCACAGCCCTGGACAGATGGCGAGAACCATG 720
    |||||
Db 661 AGAAGATGACAGCAGCTGGAAACAGATGCTCACAGCCCTGGACAGATGGCGAGAACCATG 720
Oy 721 TGAATGAGCTGGCGGGCTCTTGTACAGCAATGAGATACGTGCAAGAACGACTGACTG 780
    |||||
Db 721 TGAATGAGCTGGCGGGCTCTTGTACAGCAATGAGATACGTGCAAGAACGACTGACTG 780
Oy 781 AAGAGCTGGCTGACTGGAAGAGCGGCGGAGAGATCGCTGATCGAGGCGCCCTCCAA 840
    |||||
Db 781 AAGAGCTGGCTGACTGGAAGAGCGGCGGAGAGATCGCTGATCGAGGCGCCCTCCAA 840
Oy 841 TCTGCTGGAGACCGTCTGGAAAACTGGATTAACCTTATTAAGCAATCTCAACTTACAG 900
    |||||
Db 841 TCTGCTGGAGACCGTCTGGAAAACTGGATTAACCTTATTAAGCAATCTCAACTTACAG 900
Oy 901 GCCAACAATTAAGAACTGGAGAGCTGCGAGCAAGAAAGTCTCTACAAAGGCGACCTTA 960
    |||||
Db 901 GCCAACAATTAAGAACTGGAGAGCTGCGAGCAAGAAAGTCTCTCTACAAAGGCGACCTTA 960
Oy 961 TCTGTCAGACCGGCGCATGCTGAGAGAGAGATCGTGAGACTGTTCAAACTTAATGA 1020
    |||||
Db 961 TCTGTCAGACCGGCGCATGCTGAGAGAGAGATCGTGAGACTGTTCAAACTTAATGA 1020
Oy 1021 AGAGTGCCTTGTGGTGGAGGCGGAGCGCCCTGATGCCATGACCGGAGCGGCGCTTAG 1080
    |||||
Db 1021 AGAGTGCCTTGTGGTGGAGGCGGAGCGCCCTGATGCCATGACCGGAGCGGCGCTTAG 1080
Oy 1081 TCATCAAGACTGTGTCCAGTTCACACAGAAAGTCAAGTCTGTGTCMAATTTCTCGAGT 1140
    |||||
Db 1081 TCATCAAGACTGTGTCCAGTTCACACAGAAAGTCAAGTCTGTGTCMAATTTCTCGAGT 1140
Oy 1141 TGAATTAACACTTAAATTAAGTGTGATTAAGTAAAGACTGTGGGAGATGTGTGCC 1200
    |||||
Db 1141 TGAATTAACACTTAAATTAAGTGTGATTAAGTAAAGACTGTGGGAGATGTGTGCC 1200
Oy 1201 TCAGAGGGTCTCGGAAATTTAACTTCTGGGACAGAACACAAAAGTGTATGAACATGAGG 1260
    |||||
Db 1201 TCAGAGGGTCTCGGAAATTTAACTTCTGGGACAGAACACAAAAGTGTATGAACATGAGG 1260
Oy 1261 AGTCTAACACAGGAGCTGTCTGAGAGTTCAGAGACTTACACTTAGAGAGCAGAGAT 1320
    |||||
Db 1261 AGTCTAACACAGGAGCTGTCTGAGAGTTCAGAGACTTACACTTAGAGAGCAGAGAT 1320
Oy 1321 GTGGGAATGAGGCGCGTGTGATGCTCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1380
    |||||
Db 1321 GTGGGAATGAGGCGCGTGTGATGCTCTGATGCTCTGATGCTGATGCTGATGCTGATGCTG 1380
Oy 1381 TGAATCACTTGAAGCTGAGTGTACACCAAGGCTTAAGATTTGAGACTAGAGACCCACT 1440
    |||||
Db 1381 TGAATCACTTGAAGCTGAGTGTACACCAAGGCTTAAGATTTGAGACTAGAGACCCACT 1440

Oy 1441 CTTTCCAGTTGTGTATCTCCAAACATCTGTGAGATGCCCCAAATGCTTTGGGCAATACC 1500
    |||||
Db 1441 CTTTCCAGTTGTGTATCTCCAAACATCTGTGAGATGCCCCAAATGCTTTGGGCAATACC 1500
Oy 1501 TGTGATTAACATGCTGACCAATTAACCCCAAGAAAGTGAACCTTGTACTTAAGCCGCAA 1560
    |||||
Db 1501 TGTGATTAACATGCTGACCAATTAACCCCAAGAAAGTGAACCTTGTACTTAAGCCGCAA 1560
Oy 1561 TTTGAACCTGGGACCAAGTGGCGGAGTGTCTACGTGGAGTTCTCTCCACACCAAGC 1620
    |||||
Db 1561 TTTGAACCTGGGACCAAGTGGCGGAGTGTCTACGTGGAGTTCTCTCCACACCAAGC 1620
Oy 1621 GAGGCTGAGCATGACAGCTGACCAAGCTGGCTGGAAGCTCTGAGGCTGTGATGA 1680
    |||||
Db 1621 GAGGCTGAGCATGACAGCTGACCAAGCTGGCTGGAAGCTCTGAGGCTGTGATGA 1680
Oy 1681 ACTACTCAGGCTGCAGATCAATGAGGCTTAATCTGCAAGAAAGAAACATGGCTGGCAAG 1740
    |||||
Db 1681 ACTACTCAGGCTGCAGATCAATGAGGCTTAATCTGCAAGAAAGAAACATGGCTGGCAAG 1740
Oy 1741 GCTTCTCTTCTGGGTCTGCTAGACAAATATCATGACCTTGTGAAAAAGTATATCTTGG 1800
    |||||
Db 1741 GCTTCTCTTCTGGGTCTGCTAGACAAATATCATGACCTTGTGAAAAAGTATATCTTGG 1800
Oy 1801 CCCCTTGAATGAAGGCTACATCATGGGTTTCATCAGCAAGAGCGGGAGCGGCCATTC 1860
    |||||
Db 1801 CCCCTTGAATGAAGGCTACATCATGGGTTTCATCAGCAAGAGCGGGAGCGGCCATTC 1860
Oy 1861 TAAGCACAAAGCCCCCGGGACCTTCTACTGCGCTTACAGCAAGAGCGGAGCGGAGAG 1920
    |||||
Db 1861 TAAGCACAAAGCCCCCGGGACCTTCTACTGCGCTTACAGCAAGAGCGGAGCGGAGAG 1920
Oy 1921 GGGTCACTTCTGCTGGTGGAAAAAGACATCATGTGGCAACCAAGTCCAGTCTGAG 1980
    |||||
Db 1921 GGGTCACTTCTGCTGGTGGAAAAAGACATCATGTGGCAACCAAGTCCAGTCTGAG 1980
Oy 1981 AGCCATACCAACAGCAGCTGAAACAATGATTTGCTGAATCATCATGTGGCTATA 2040
    |||||
Db 1981 AGCCATACCAACAGCAGCTGAAACAATGATTTGCTGAATCATCATGTGGCTATA 2040
Oy 2041 AGATCATGATGCGACCAACATCTGCTGCTCCACTTGTCTACCTTACCCCGACATTC 2100
    |||||
Db 2041 AGATCATGATGCGACCAACATCTGCTGCTCCACTTGTCTACCTTACCCCGACATTC 2100
Oy 2101 CCAAGAGAGAGCAATTGGAAAGTACTGTAGGCCGAGAGCCAGAGACACCCGAAACCG 2160
    |||||
Db 2101 CCAAGAGAGAGCAATTGGAAAGTACTGTAGGCCGAGAGCCAGAGACACCCGAAACCG 2160
Oy 2161 ACCAGTAGTGTGCCCCGCTACCTGAAGAGCAAGTTCTATGTGTGACACCAAGACT 2220
    |||||
Db 2161 ACCAGTAGTGTGCCCCGCTACCTGAAGAGCAAGTTCTATGTGTGACACCAAGACT 2220
Oy 2221 GCAGCAATACCATTTGACCTGCGATGTCGCCGACCTTGTGATTCATTTGATGAGAGTTT 2280
    |||||
Db 2221 GCAGCAATACCATTTGACCTGCGATGTCGCCGACCTTGTGATTCATTTGATGAGAGTTT 2280
Oy 2281 GAAATTAACGTGAAGTGTCTAGCCCTCAGCAGAGAGGCAAGTTGAGTGCCTCAGGTTTG 2340
    |||||
Db 2281 GAAATTAACGTGAAGTGTCTAGCCCTCAGCAGAGAGGCAAGTTGAGTGCCTCAGGTTTG 2340
Oy 2341 ACATGATCTGACCTCGGAGTGTCTACCTTCCCATGTGAGAGCTGAAMACGAAAGCT 2400
    |||||
Db 2341 ACATGATCTGACCTCGGAGTGTCTACCTTCCCATGTGAGAGCTGAAMACGAAAGCT 2400
Oy 2401 GCAGAGACTGACTTGAACACCTGCCCCGCTGTCACACCCCTAAGCAGACGCAACCCATA 2460
    |||||
Db 2401 GCAGAGACTGACTTGAACACCTGCCCCGCTGTCACACCCCTAAGCAGACGCAACCCATA 2460
Oy 2461 TCGTCTGAACCTCTTAACCTTGTGTGCTCAGATTTTCTTATTTCTACTTCTGCT 2520
    |||||
Db 2461 TCGTCTGAACCTCTTAACCTTGTGTGCTCAGATTTTCTTATTTCTACTTCTGCT 2520
```

OY 2521 ATCTTTGGCAATCTGGGCACTTTTAAAAAGAGAAATGAGTGTGGTGTGATTAAC 2580
|||||
Db 2521 ATCTTTGGGCAATCTGGGCACTTTTAAAAAGAGAAATGAGTGTGGTGTGATTAAC 2580
OY 2581 TGTATGTAAAGAGAGAGACTCTGTAGTGTGGGATGGGGCTGAGAGCAGAGAGGAGC 2640
|||||
Db 2581 TGTATGTAAAGAGAGAGACTCTGTAGTGTGGGATGGGGCTGAGAGCAGAGAGGAGC 2640
OY 2641 AAAGGGAAACACCTCTCTGCTCCGCGGCTCCCTCTTTTACAGACGCTCGGGGCTTG 2700
|||||
Db 2641 AAAGGGAAACACCTCTCTGCTCCGCGGCTCCCTCTTTTACAGACGCTCGGGGCTTG 2700
OY 2701 TTGTGTAGACAAGTGGCCCTGTGGTCCCATGCTACCTGTGGCCCACTGTGTAGCTGAT 2760
|||||
Db 2701 TTGTGTAGACAAGTGGCCCTGTGGTCCCATGCTACCTGTGGCCCACTGTGTAGCTGAT 2760
OY 2761 ACCGCATTTGGGAACTCTGGCTCTGCTACCTTTTCACCTTGTCTAATATCCACATAGAGC 2820
|||||
Db 2761 ACCGCATTTGGGAACTCTGGCTCTGCTACCTTTTCACCTTGTCTAATATCCACATAGAGC 2820
OY 2821 TAGGACTAAGCCCAAGAGGTTCTCTTAATTAATTAATTAATTAATTAATTAATTAAT 2869
|||||
Db 2821 TAGGACTAAGCCCAAGAGGTTCTCTTAATTAATTAATTAATTAATTAATTAATTAAT 2869

RESULT 5
AR153599 2869 bp DNA PAT 08-AUG-2001
LOCUS AR153599 Sequence 9 from patent US 6235873.
DEFINITION AR153599
ACCESSION AR153599 GI:15121131
VERSION AR153599.1
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2869)
AUTHORS Bromberg,J.F., Wlzeszczynska,M.H., Zhao,Y. and Darnell,J.E. Jr.
TITLE Constitutively active transcription factors and their uses for
identifying modulators of activity including dysproliferative
cellular changes
JOURNAL Patent: US 6235873-A 9 22-May-2001;
FEATURES
source 1..2869 Location/Qualifiers
BASE COUNT 753 a 742 c 772 g 602 t
ORIGIN

Query Match 99.7%: Score 2861; DB 6; Length 2869;
Best Local Similarity 99.8%: Pred. No. 0;
Matches 2864: Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 GCCCGC,ACACAGCGCCGCGCAGTGGGCTCAGCCCGAGAGACAGTGCAGCCCTGACT 60
|||||
Db 1 GCCCGCAGCCAGCGCGCGGCGCAAGTGGGCTCAGCCCGAGAGACAGTGCAGCCCTGACT 60
OY 61 GCACAGAGATGGCTCAGTGAACAGCTGCAGACAGCTGGACACAGCTTGAAGCAGC 120
|||||
Db 61 GCACAGAGATGGCTCAGTGAACAGCTGCAGACAGCTGGACACAGCTTGAAGCAGC 120
OY 121 TGCCACCACTGTACAGGACAGCTGCCCATGGAGCTGGGAGTTCTCTGGACCTTGA 180
|||||
Db 121 TGCCACCACTGTACAGGACAGCTGCCCATGGAGCTGGGAGTTCTCTGGACCTTGA 180
OY 181 TTGAGAGTCAAGACTGGGCATATGCAAGCCAGCAAGAGTCAATGCCAGTTGTGTTTC 240
|||||
Db 181 TTGAGAGTCAAGACTGGGCATATGCAAGCCAGCAAGAGTCAATGCCAGTTGTGTTTC 240
OY 241 ATATCTCTTTGGGTGAATTTGACGCAATATAGCCGATTTCTGTGAAGATTCATGTC 300
|||||
Db 241 ATATCTCTTTGGGTGAATTTGACGCAATATAGCCGATTTCTGTGAAGATTCATGTC 300
OY 301 TCTATCAGCACAACCTTGAGAGAAATCAAGCAGTTTCTGACAGCAGTATCTTGAGAGC 360
|||||

|||||
Db 301 TCTATCAGCACAACCTTGAGAGAAATCAAGCAGTTTCTGACAGCAGTATCTTGAGAGC 360
OY 361 CAATGCAAAATTTGCCCGGATCTGTGGCCCGATGCTGTGGGAAAGATCTGCTCTCCACA 420
|||||
Db 361 CAATGCAAAATTTGCCCGGATCTGTGGCCCGATGCTGTGGGAAAGATCTGCTCTCCACA 420
OY 421 CGGAGCTACGGCAGCCAGCAGAGGGGCGGCAAGCCCAACCCCAAGCCGCGTACTGA 480
|||||
Db 421 CGGAGCTACGGCAGCCAGCAGAGGGGCGGCAAGCCCAACCCCAAGCCGCGTACTGA 480
OY 481 CAGAGAGCAGCAGATGTTGAGCAGCATCTTACGATGTCCGAGACGAGTGCAGATC 540
|||||
Db 481 CAGAGAGCAGCAGATGTTGAGCAGCATCTTACGATGTCCGAGACGAGTGCAGATC 540
OY 541 TAGAAGCAAAATGAAGGTGTGAGAACTCCAGCAGCACTTGTGATTTCACTACAAA 600
|||||
Db 541 TAGAAGCAAAATGAAGGTGTGAGAACTCCAGCAGCACTTGTGATTTCACTACAAA 600
OY 601 CCTCAAGAGCCCAAGAGACATGAGGATCTGAATGGAACCAACCACTGTGACAGAC 660
|||||
Db 601 CCTCAAGAGCCCAAGAGACATGAGGATCTGAATGGAACCAACCACTGTGACAGAC 660
OY 661 AGAAGATGCAGCAGCTGGAAACAGATGCTCACAGCCCTGGACAGATGCGGAGAACTTG 720
|||||
Db 661 AGAAGATGCAGCAGCTGGAAACAGATGCTCACAGCCCTGGACAGATGCGGAGAACTTG 720
OY 721 TGAGTACGCTGGCGGGCTCTTGTACAGCAATGAGTACGTCAGAGACACTGATG 780
|||||
Db 721 TGAGTACGCTGGCGGGCTCTTGTACAGCAATGAGTACGTCAGAGACACTGATG 780
OY 781 AAGAGCTGGCTGACTGGAAGAGCGGCGCAGATGCGCTGATGGAGCCCTCCACA 840
|||||
Db 781 AAGAGCTGGCTGACTGGAAGAGCGGCGCAGATGCGCTGATGGAGCCCTCCACA 840
OY 841 TCTGCTGGACCGCTGGAAACCTGSAATTAATTAAGCAATTCGAAATTCGATGAGACC 900
|||||
Db 841 TCTGCTGGACCGCTGGAAACCTGSAATTAATTAAGCAATTCGAAATTCGATGAGACC 900
OY 901 GCCAACAATTAAGAACTGAGAGAGCTGACAGAGAAAGTGTCTTACAGGGCGACCTTA 960
|||||
Db 901 GCCAACAATTAAGAACTGAGAGAGCTGACAGAGAAAGTGTCTTACAGGGCGACCTTA 960
OY 961 TCGTGCAGCACCGGCCCATGCTGAGAGAGAGATCTGTGACCTGTTCAAGAACTTAATGA 1020
|||||
Db 961 TCGTGCAGCACCGGCCCATGCTGAGAGAGAGATCTGTGACCTGTTCAAGAACTTAATGA 1020
OY 1021 AGAGTCCCTTGTGTGAGAGCGGAGCCCTGATGCCATGCCAGCCGAGCCGCTTAG 1080
|||||
Db 1021 AGAGTCCCTTGTGTGAGAGCGGAGCCCTGATGCCATGCCAGCCGAGCCGCTTAG 1080
OY 1081 TCATCAAGACTGTGTCCAGTTTACACGAAAGTCAAGTTGCTGCTCAATTTCTGAGT 1140
|||||
Db 1081 TCATCAAGACTGTGTCCAGTTTACACGAAAGTCAAGTTGCTGCTCAATTTCTGAGT 1140
OY 1141 TGAATTAACACTTAAATTAAGTGTGATTAAGTCAATTAAGTGTGAGTGTGCTGCC 1200
|||||
Db 1141 TGAATTAACACTTAAATTAAGTGTGATTAAGTCAATTAAGTGTGAGTGTGCTGCC 1200
OY 1201 TCAGAGGCTCTCGGAAATTAATTAAGTGTGAGCAGCAACCAAAAGTGAATGAGAG 1260
|||||
Db 1201 TCAGAGGCTCTCGGAAATTAATTAAGTGTGAGCAGCAACCAAAAGTGAATGAGAG 1260
OY 1261 AGTCTAACAGCGCAGCTGTCTGACAGTTTCAACACCTTGTGAGAGAGAGAT 1320
|||||
Db 1261 AGTCTAACAGCGCAGCTGTCTGACAGTTTCAACACCTTGTGAGAGAGAGAT 1320
OY 1321 GTGGGAATGAGAGCGGCTGCAATGTGATGCTCTTGTGATGCTGAGAGAGCTGAC 1380
|||||
Db 1321 GTGGGAATGAGAGCGGCTGCAATGTGATGCTCTTGTGATGCTGAGAGAGCTGAC 1380
OY 1381 TGATCACCTTGAGAGCTGAGGTGTACCAAGGCTCAAGATTAAGCTAGAGACCACT 1440
|||||

Db 1381 TGATCACCCTTCGACATGAGCTGTACCAAGCCCTCAAGATTGACCTAGACAGCCACT 1440
 Oy 1441 CCTTGCAGTTGGTGTGATCTCCAAACATCTGTGATGCCAAATGCTTGGGCATCAATCC 1500
 Db 1441 CCTTGCAGTTGGTGTGATCTCCAAACATCTGTGATGCCAAATGCTTGGGCATCAATCC 1500
 Oy 1501 TGTGCTATTAACATGCTCAACCAATTAACCCCAAGAACGTGAACCTTTCTACTAAGCCGCCAA 1560
 Db 1501 TGTGCTATTAACATGCTCAACCAATTAACCCCAAGAACGTGAACCTTTCTACTAAGCCGCCAA 1560
 Oy 1561 TTGGAACCTGGAGCAAGTGGCCGAGGTGCTCAGCTGGAGTTCTGCTCCACACCAAGC 1620
 Db 1561 TTGGAACCTGGAGCAAGTGGCCGAGGTGCTCAGCTGGAGTTCTGCTCCACACCAAGC 1620
 Oy 1621 TTTGAACCTGGAGCAAGTGGCCGAGGTGCTCAGCTGGAGTTCTGCTCCACACCAAGC 1680
 Db 1621 TTTGAACCTGGAGCAAGTGGCCGAGGTGCTCAGCTGGAGTTCTGCTCCACACCAAGC 1680
 Oy 1681 ACTACTCAGGGTGTCAATGATCAGTGGGCTAAATCTGCAAGAAAGAACATGGCTGCCAAG 1740
 Db 1681 ACTACTCAGGGTGTCAATGATCAGTGGGCTAAATCTGCAAGAAAGAACATGGCTGCCAAG 1740
 Oy 1741 GCTTCTCCTTCTGCGTCTGCTAGACAAATATCATGACCTTGTGAAAAAGTATATCTTGG 1800
 Db 1741 GCTTCTCCTTCTGCGTCTGCTAGACAAATATCATGACCTTGTGAAAAAGTATATCTTGG 1800
 Oy 1801 CCCCTTGGAAATGAAGGGTACATCGGTTTCATCAGCAAGAGGAGGCGGCATCC 1860
 Db 1801 CCCCTTGGAAATGAAGGGTACATCGGTTTCATCAGCAAGAGGAGGCGGCATCC 1860
 Oy 1861 TTAAGCAAAAGCCCCCGGACCTCTCTACTGCGCTTACGAGAGAGCAAGAAAGAGG 1920
 Db 1861 TTAAGCAAAAGCCCCCGGACCTCTCTACTGCGCTTACGAGAGAGCAAGAAAGAGG 1920
 Oy 1921 GGGTCACTTCTGCTGGTGGTGAAGGACATCACTGGCAACCCAGATCCAGTCTGTAG 1980
 Db 1921 GGGTCACTTCTGCTGGTGGTGAAGGACATCACTGGCAACCCAGATCCAGTCTGTAG 1980
 Oy 1981 AGCCATACACCAAGCAGAGCTGAACAACATGATTTGCTGAATCATCATGGGCTATA 2040
 Db 1981 AGCCATACACCAAGCAGAGCTGAACAACATGATTTGCTGAATCATCATGGGCTATA 2040
 Oy 2041 AGATCATGATGCGACCAACATCTGCTGCTCCACTTGTACTTACCTTACCCGACATTC 2100
 Db 2041 AGATCATGATGCGACCAACATCTGCTGCTCCACTTGTACTTACCTTACCCGACATTC 2100
 Oy 2101 CCAAGGAGAGGCAATTTGGAAGTACTGTAGGCCGAGAGCAGAGACCCCGAAGCG 2160
 Db 2101 CCAAGGAGAGGCAATTTGGAAGTACTGTAGGCCGAGAGCAGAGACCCCGAAGCG 2160
 Oy 2161 ACCGAGCTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
 Db 2161 ACCGAGCTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
 Oy 2221 GCAGCAATACCATTTGACCTGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
 Db 2221 GCAGCAATACCATTTGACCTGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
 Oy 2281 GAAATTAACGGTGAAGGTGCTGAGCCCTCAGCAGAGGAGGAGTTGAGTGGCTCAAGTTTG 2340
 Db 2281 GAAATTAACGGTGAAGGTGCTGAGCCCTCAGCAGAGGAGGAGTTGAGTGGCTCAAGTTTG 2340
 Oy 2341 ACATGATCTGACCTGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
 Db 2341 ACATGATCTGACCTGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
 Oy 2401 GCAGAGAGTACTTGAACACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
 Db 2401 GCAGAGAGTACTTGAACACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
 Oy 2461 TCGTCTGAACCTCTAATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
 Db 2461 TCGTCTGAACCTCTAATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520

Oy 2521 ATCTTTGGGCAATCTGGGCACTTTTAAAGAGAGAAATGAGTGTGTGGTATTAAC 2580
 Db 2521 ATCTTTGGGCAATCTGGGCACTTTTAAAGAGAGAAATGAGTGTGTGGTATTAAC 2580
 Oy 2581 TGTATGTAAAGAGAGAGACCTGTGAGTGTGGGATGTGGGCTGAGACAGAAAGAGAGC 2640
 Db 2581 TGTATGTAAAGAGAGAGACCTGTGAGTGTGGGATGTGGGCTGAGACAGAAAGAGAGC 2640
 Oy 2641 AAAGGGAGACACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
 Db 2641 AAAGGGAGACACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
 Oy 2701 TTTTAAACAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
 Db 2701 TTTTAAACAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
 Oy 2761 ACCCATTTCTGGAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
 Db 2761 ACCCATTTCTGGAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
 Oy 2821 TAGACTTAAGCCGAGAGGTTCTTTAATTAAAAAAGAGAGAGAGAGAGAGAGAGAGAG 2869
 Db 2821 TAGACTTAAGCCGAGAGGTTCTTTAATTAAAAAAGAGAGAGAGAGAGAGAGAGAGAG 2869
 RESULT 6
 LOCUS BC003806 2964 bp mRNA ROD 12-JUL-2001
 DEFINITION Mus musculus, Signal transducer and activator of transcription 3, clone MGC:6089 IMAGE:359358, mRNA, complete cds.
 ACCESSION BC003806
 VERSION BC003806.1 GI:13277851
 KEYWORDS MGC.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2964)
 REFERENCE Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (28-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 COMMENT Contact: MGC help desk
 Email: gcabs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) mdgpaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
 FEATURES
 source
 1..2964
 /organism="Mus musculus"
 /db_xref="locusID:25125"
 /db_xref="taxon:10090"
 /clone="MGC:6089 IMAGE:359358"
 /tissue_type="Mammary tumor, Metallothionien-TGF alpha model, 10 month old virgin mouse, Taken by biopsy."
 /clone_id="NCI_CGAP_Mam1"
 /lab_host="DH10b"

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA 91: 476/15.

CDS

/note="Vector: PCMV-SPORT6"
168..2480
/codon_start=1
/product="Signal transducer and activator of transcription
3"
/protein_id="AAH03806.1"
/db_xref="GI:13277852"
/translation="MAQMNQLOLDTRYLEQHLQYLSDFPMELEQFLAPWESQMA
YAASKESHATLVEFHNILGELIDQYSLFLOSNNLYQNLNRIKQFLDSRYLEKMEIA
RIVARCLMEESRLQTAATAOOGQANHTAAVTEKQMLQEOHLODVRKRVODLEQ
KMKVVENIADDPEFNYKTLKSQGMQDLNNOSVTRQKQKQLEQMLTALDOMRSIV
SELAGLISAMEYOKTLIDELADMKRRQOACIGSPNIDLELNMITSLSAQSD
TROIKKLEAMEYOKLSYKQDPIVQHRPMLEERLELFRNLMSAFVVERQPCMHND
RPLVKTGVQFTTKVRLLVKFPPELNTOLKIKVCIDKRSQVYALRGSRNIIJSTNK
VMNEESNNSLSAEFKHLLREQRCNGSNGRANCDASLIVTEHLHLTFTEVYHQGL
KIDLETHSLPVVVISNICOMPNAWASILWMLNMLNPNKNVNEFKPIGTMDOVAEYL
SWQPSYTKRGLSLEQLTLAEKLIGPVNNSGCOITMAKFCXKNMAGKGSFEMVMD
NIIDLVKYLIALNVEGYIMGFISKEERAILSTKPPGTFLLRESSEKSGEYTFVY
EKDISGKTOIQSVEPYTKKQOLNNMSFAEIIIMGYIMQIMATNIVLSPVLYLPDIKKEA
FGKYCRPESQEHPEPDSAPYLKTKPFICTYPTTCSNTIDLPHSPRLDSLMQFGNN
GEGAEPSAGGFESLTFMDLTSFCATSPM"

BASE COUNT 764 a 782 c 807 g 611 t
ORIGIN

Query Match 96.5%; Score 2767.8; DB 10; Length: 2964;

Best Local Similarity 99.2%; Pred. No. 0; Mismatches 17; Indels 7; Gaps 5;

Matches 2835; Conservative 0; Mismatches 17; Indels 7; Gaps 5;

QY 13 CAGCCGCGCAGTCCGGC--TCAGCCCGAGACAGTCGAGACCCG1JACTGCGAGCAGAT 70
DB 110 CAGCCGCGCAGTCCGGCAGCTCGAGCAGACGACGACCCCTGACTGCGAGCAGAT 169
QY 71 GGCTAGTGAACCGCTGCGACGACGCTACGCTACGACGACGCTGACGACGCT 130
DB 170 GGCTAGTGAACCGCTGCGACGACGCTACGCTACGACGACGCTGACGACGCT 229
QY 131 GTACAGCGACAGTTCGCCATGAGCTGGGCGAGTTCTGGCAGCTTGGATTGAGATCA 190
DB 230 GTACAGCGACAGTTCGCCATGAGCTGGGCGAGTTCTGGCAGCTTGGATTGAGATCA 289
QY 191 ACACGCGCATATGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 250
DB 290 ACACGCGCATATGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 349
QY 251 GGGTAAATTTGACGACATATAGCCGATTCCTGCAAGAGTCGAATGTCCTATACGA 310
DB 350 GGGTAAATTTGACGACATATAGCCGATTCCTGCAAGAGTCGAATGTCCTATACGA 409
QY 311 CAACCTTGACAGATCAGCAATGTTCTGACAGCAGGTATCTTGAAGAGCCAAATGAAT 370
DB 410 CAACCTTGACAGATCAGCAATGTTCTGACAGCAGGTATCTTGAAGAGCCAAATGAAT 469
QY 371 TCCCGGATCGTGGCCGATGCTGCGGAGAGTCTGCTGCTCCAGACGCGACGAC 430
DB 470 TCCCGGATCGTGGCCGATGCTGCGGAGAGTCTGCTGCTCCAGACGCGACGAC 529
QY 431 GGCAGCCCGACAGAGGGGCGCAGGCCAACCCAAAGCCCGCTGATGACAGAGCA 490
DB 530 GGCAGCCCGACAGAGGGGCGCAGGCCAACCCAAAGCCCGCTGATGACAGAGCA 589
QY 491 GGAGATGTTGGAGCAGCATCTTCAGAGATTCGGAAGCAGTGCAGAGATCTGAAGAGAA 550
DB 590 GGAGATGTTGGAGCAGCATCTTCAGAGATTCGGAAGCAGTGCAGAGATCTGAAGAGAA 649
QY 551 AATGAAGTGTGAGAGAACCTTCAGAGACGCTTTGATTTCACTACAAAACCTCAAGAG 610
DB 650 AATGAAGTGTGAGAGAACCTTCAGAGACGCTTTGATTTCACTACAAAACCTCAAGAG 709
QY 611 CCAAGGAGATGACGAGATCTGAATGGAACAACAGCTGTGACGACAGCAAGATGCA 670
DB 710 CCAAGGAGATGACGAGATCTGAATGGAACAACAGCTGTGACGACAGCAAGATGCA 769

QY 671 GCAGCTGAACAGATGCTCAAGCCCTGGACCAAGATGGGAGAGCATTTGATGAGCT 730
DB 770 GCAGCTGAACAGATGCTCAAGCCCTGGACCAAGATGGGAGAGCATTTGATGAGCT 829
QY 731 GCGCGGCTCTTGTACGACATGAGTACGTGACAGAGACATGATGAGAGCTGCG 790
DB 830 GCGCGGCTCTTGTACGACATGAGTACGTGACAGAGACATGATGAGAGCTGCG 889
QY 791 TGACTGGAAGAGCGCGCGCAGAGATCGCGTGCATGGAGAGCCCTCCCAACATCTCCCGGA 850
DB 890 TGACTGGAAGAGCGCGCGCAGAGATCGCGTGCATGGAGAGCCCTCCCAACATCTCCCGGA 949
QY 851 CCGTGTGAAAACCTGATTAATCTTCAATTCATGCAATTCATGACACCCCAACAAAT 910
DB 950 CCGTGTGAAAACCTGATTAATCTTCAATTCATGCAATTCATGACACCCCAACAAAT 1009
QY 911 TAAGAACTGGAGAGAGCTGACAGCAGAAAGTGTCTACAAAGGCGACCTATCGTCAGCA 970
DB 1010 TAAGAACTGGAGAGAGCTGACAGCAGAAAGTGTCTACAAAGGCGACCTATCGTCAGCA 1069
QY 971 CCGGCGCATGCTGGAGGAGAGATCGTGAGCTGTTCAGAACTTAATGAAGATGCTT 1030
DB 1070 CCGGCGCATGCTGGAGGAGAGATCGTGAGCTGTTCAGAACTTAATGAAGATGCTT 1129
QY 1031 CCGTGTGAGAGCGCGCAGCCCTGCATGACCCGAGACCGGCGCTTATGTCATCAAGAC 1090
DB 1130 CCGTGTGAGAGCGCGCAGCCCTGCATGACCCGAGACCGGCGCTTATGTCATCAAGAC 1189
QY 1091 TCGTGTCCAGTTTACCAGCAAGTACGTTGCTGTCAAATTTCTGTGATGAATATCA 1150
DB 1190 TCGTGTCCAGTTTACCAGCAAGTACGTTGCTGTCAAATTTCTGTGATGAATATCA 1249
QY 1151 GCTTAAATTTAAGTGTGATTAAGACTCTGGGATTTGCTGCGCTCAGAGGCTC 1210
DB 1250 GCTTAAATTTAAGTGTGATTAAGACTCTGGGATTTGCTGCGCTCAGAGGCTC 1309
QY 1211 TCGAAATTTAATCAATTTGCGCAGCAGACAAAGATGATGAACATGAGAGCTTAACAA 1270
DB 1310 TCGAAATTTAATCAATTTGCGCAGCAGACAAAGATGATGAACATGAGAGCTTAACAA 1369
QY 1271 GGGCAACCTGTCGAGAGTTCAACGACCTGACCTTGGGAGAGAGATGTGGAAATGG 1330
DB 1370 GGGCAACCTGTCGAGAGTTCAACGACCTGACCTTGGGAGAGAGATGTGGAAATGG 1429
QY 1331 AGCGCGTCCAAATTTGATGCTTCCCTTGAATGCTGACTGAGAGCTGACCTGATCACCTT 1390
DB 1430 AGCGCGTCCAAATTTGATGCTTCCCTTGAATGCTGACTGAGAGCTGACCTGATCACCTT 1489
QY 1391 CGAGACTGAGGTGTACGACCAAGGCTCAAGATTGACCTAGAGACCACTCCTTGCCAGT 1450
DB 1490 CGAGACTGAGGTGTACGACCAAGGCTCAAGATTGACCTAGAGACCACTCCTTGCCAGT 1540
QY 1451 TGTGTGATCTTCAACATCTGTCAAGATGCAATGCTTGGGATCAATCTGTGGTATTA 1510
DB 1550 TGTGTGATCTTCAACATCTGTCAAGATGCAATGCTTGGGATCAATCTGTGGTATTA 1609
QY 1511 CATGCTGACCAATTAACCCCAAGACGTAACTTTTCACTAAAGCCGCAATTTGAACCTG 1570
DB 1610 CATGCTGACCAATTAACCCCAAGACGTAACTTTTCACTAAAGCCGCAATTTGAACCTG 1669
QY 1571 GGACCAAGTGGCGGAGGTGCTAGCTGGAGTTCGATGTCACACCAAGCAAGCAAGGCTGAG 1630
DB 1670 GGACCAAGTGGCGGAGGTGCTAGCTGGAGTTCGATGTCACACCAAGCAAGCAAGGCTGAG 1729
QY 1631 CATGAGACGCTGACACAGCGTGGCTGAGAGAGCTCTAGGCGCTGTGTGAATCTACAGG 1690
DB 1730 CATGAGACGCTGACACAGCGTGGCTGAGAGAGCTCTAGGCGCTGTGTGAATCTACAGG 1789
QY 1691 GTGTGAGATCACATGGGCTAAATTTCTGCAAGAAAGAACTGCTGCAAGAGGCTTCTCTT 1750
DB 1790 GTGTGAGATCACATGGGCTAAATTTCTGCAAGAAAGAACTGCTGCAAGAGGCTTCTCTT 1849
QY 1751 CTGGGCTGTGCTAGCAATATATCATGACCTTGTGAAAAAATATATCTTGGCAA 1810

Db	235	GTACAGCACACGTTCCCATATGAGACTGCGGACGATTCCTCGGCACCTTGGATTGAGAGTCA	294
OY	191	AGACTGGGCATATGCAAGCCAGCAAGAGTGCACATGCCAGTTGGTGTTCATPAATCTTT	250
Db	295	AGACTGGGATATGCAAGCCAGCAAGAGTGCACATGCCAGTTGGTGTTCATPAATCTTT	354
OY	251	GGGTCAATTTGACAGCAATATAGCCGATTCCTGCAGAGTGCCAAATGTCCTCATAGCA	310
Db	355	GGGTCAATTTGACAGCAATATAGCCGATTCCTGCAGAGTGCCAAATGTCCTCATAGCA	414
OY	311	CAACCTTGCAGAAATCAACACGATTTCTGCAGAGCAGATATCTTGAAGACCAATGGAAAT	370
Db	415	CAACCTTGCAGAAATCAACACGATTTCTGCAGAGCAGATATCTTGAAGACCAATGGAAAT	474
OY	371	TGCCCCGATGTGCCCCGATGCTGTTGGAGAGTCTGCCCTCTCTCAGAGCCGACCAC	430
Db	475	TGCCCGATGTGCCCCGATGCTGTTGGAGAGTCTGCCCTCTCTCAGAGCCGACCAC	534
OY	431	GGCAGCCGACGAAGGGGGGCGAGGCCAACCAACCACAGCGCGCGTAGTGACAGAGACA	490
Db	535	GGCAGCCGACGAAGGGGGGCGAGGCCAACCAACCACAGCGCGCGTAGTGACAGAGACA	594
OY	491	GCAGATGTTGAGACAGCAGATCTTCAGGATGTCCGGAAGCAGTGCAGAGATCTAGAACAA	550
Db	595	GCAGATGTTGAGACAGCAGATCTTCAGGATGTCCGGAAGCAGTGCAGAGATCTAGAACAA	654
OY	551	AATGAAGTGTGAGACACCTCCAGACGACTTGAATTTCACTACAAAACCCCTCAAGAG	610
Db	655	AATGAAGTGTGAGACACCTCCAGACGACTTGAATTTCACTACAAAACCCCTCAAGAG	714
OY	611	CCAAAGAGACATCAGAGATCTGAATGGAACCAACAGTCTGTGACAGACAGAGAGATGCA	670
Db	715	CCAAAGAGACATCAGAGATCTGAATGGAACCAACAGTCTGTGACAGACAGAGAGATGCA	774
OY	671	GCAGCTGGAACACATGCTCACAGCCCTGGACCAAGATCCGAGAAGCTTGGAGTGAAGCT	730
Db	775	GCAGCTGGAACACATGCTCACAGCCCTGGACCAAGATCCGAGAAGCTTGGAGTGAAGCT	834
OY	731	GGCGGGGCTTGTTCAGCAATGGAGTACGTGCAGAAAGACACTGATGTAAGAGCTGGC	790
Db	835	GGCGGGGCTTGTTCAGCAATGGAGTACGTGCAGAAAGACACTGATGTAAGAGCTGGC	894
OY	791	TGACTGGAAAGCGCGCCAGAGATCCGCTGCTATCGAGGGCCCTCCCAACATCTGCCCTGA	850
Db	895	TGACTGGAAAGCGCGCGAGAGATCCGCTGCTATCGAGGGCCCTCCCAACATCTGCCCTGA	954
OY	851	CCCTCTGGAAGAACTGATTAATCTTATAGCAGATCTCAACTTCAGACCCGCCCAACAT	910
Db	955	CCCTCTGGAAGAACTGATTAATCTTATAGCAGATCTCAACTTCAGACCCGCCCAACAT	1014
OY	911	TAAAGAACTGGAGAGCTGCAGCAGAAAGTCTCTACAAGGGCGACCTATCTGTGCAGCA	970
Db	1015	TAAAGAACTGGAGAGCTGCAGCAGAAAGTCTCTACAAGGGCGACCTATCTGTGCAGCA	1074
OY	971	CCGGCCCATCTGGAGAGAGGATCGTAGAGCTGTGCAGAACTTAATGAAGATGCTCT	1030
Db	1075	CCGGCCCATCTGGAGAGAGGATCGTAGAGCTGTGCAGAACTTAATGAAGATGCTCT	1134
OY	1031	CGTGCTGAGACCGAGCCCTGCATGCCATGCACCCGGACCGCCCTTAGTCAATCAAGAC	1090
Db	1135	CGTGCTGAGACCGAGCCCTGCATGCCATGCACCCGGACCGCCCTTAGTCAATCAAGAC	1194
OY	1091	TGGTGTCACTTTACCGAAGAAAGTCAAGTTGCTGTCAAAATTTCTGTGATTGAATTATCA	1150
Db	1195	TGGTGTCACTTTACCGAAGAAAGTCAAGTTGCTGTCAAAATTTCTGTGATTGAATTATCA	1254
OY	1151	GCTTAAATTTAAAGTGTGATTTGATTAAGACTCTGGGGATTTGCTGCCCTCAGAGGGTTC	1210
Db	1255	GCTTAAATTTAAAGTGTGATTTGATTAAGACTCTGGGGATTTGCTGCCCTCAGAGGGTTC	1314
OY	1211	TGCGAAATTTAACTTCTGGGACAGCAACAAAAAGTATGAACAATGGAGAGTCTCAACA	1270
Db	1315	TGCGAAATTTAACTTCTGGGACAGCAACAAAAAGTATGAACAATGGAGAGTCTCAACA	1374

QY	2351	GACCCGAGAGTGTCTACTCTCCCTCCCATGTGAGGAGCTAAACACGAAACTGTCAGACAGT	2410
Db	2405	GACCTCGGAGTGTGTACTCTCCCTCCCATGTGAGAGCTTAACACGAAACTGTCAGACAGT	2464
QY	2411	GACTTGAGACACTCTGCCCGCTGTCCACCCCTTAACGACGCGAACCCTATCTGTGAAA	2470
Db	2465	GACTTGAGACACTCTGCCCGCTGTCCACCCCTTAAGACGCGAACCCTATCTGTGAAA	2524
QY	2471	CTCCCTAACTTTGTGGTTCGACATTTTTTTTTTAATTCCTACCTCTGCTATCTTTGGGC	2530
Db	2525	CTCCCTAACTTTGTGGTTCGACATTTTTTTTTTAATTCCTACCTCTGCTATCTTTGGGC	2584
QY	2531	AATCTGGGCACTTTTTTAAAGAGAGAAATGAGTGAATGGGTGATTAACCTGTATGTAA	2590
Db	2585	AATCTGGGCACTTTTTTAAAGAGAGAAATGAGTGAATGGGTGATTAACCTGTATGTAA	2644
QY	2591	AGAGGAGAG - ACCTGTGAGTCTGGGGAGTGGGGCTAGACGACGAGGAGGCGCAAGGGGAA	2649
Db	2645	AGAGGAGAGCACTCTGTGAGTCTGGGGAGTGGGGCTAGAGCAGAAAGGAGGCGCAAGGGGAA	2704
QY	2650	CACCTCTGTCTCCCTGGCCGCTCCCTCTCTTTTCAGCAGCTCGGGGTTGTGTGTAGAC	2709
Db	2705	CACCTCTGTCTCCCTGGCCGCTCCCTCTCTTTTCAGCAGCTCGGGGTTGTGTGTAGAC	2764
QY	2710	AAGTGCCTCTGTGTGGTCCCATGGCTACCTGTGTGGCCCACTCTGTGAGCTGATACCCATTG	2769
Db	2765	AAGTGCCTCTGTGTGGTCCCATGGCTACCTGTGTGGCCCACTCTGTGAGCTGATACCCATTG	2824
QY	2770	TGGGAACCTCGTGGCTGTGACACTTTAACCTTGCTAAATTCGACATAGAAAGCTAGACATA	2829
Db	2825	TGGGAACCTCGTGGCTGTGACACTTTAACCTTGCTAAATTCGACATAGAAAGCTAGACATA	2884
QY	2830	GCCGAGAGAGTTCTCTTTAAATTAACCAAAAAAAAAAAAAAAAAA	2869
Db	2885	GCCGAGAGAGTTCTCTTTAAATTAACCAAAAAAAAAAAAAAAAAA	2924
RESULT	8		
LOCUS	MMU08378		
DEFINITION	Mus musculus acute phase response factor (APRF) mRNA, complete cds.		
VERSION	U08378		
KEYWORDS	U08378.1 GI:473889		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 2924)		
TITLE	Raz,R., Durbin,J.E. and Levy,D.E.		
JOURNAL	Acute phase response factor and additional members of the		
MEDLINE	Interleukin-stimulated gene factor 3 family integrate diverse		
REFERENCE	signals from cytokines, interferons, and growth factors		
AUTHORS	J. Biol. Chem. 269, 24391-24395 (1994)		
TITLE	2 (bases 1 to 2924)		
JOURNAL	Levy,D.E.		
MEDLINE	Direct Submission		
REFERENCE	Submitted (05-APR-1994) David E. Levy, Pathology, New York		
AUTHORS	University School of Medicine, 550 First Avenue, New York, NY		
TITLE	10016, USA		
FEATURES	Location/Qualifiers		
SOURCE	1..2924		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	/clone="pBS30"		
	/clone_lib="Unizap library (Stratagene)"		
	/tissue_type="brain"		
	/dev_stage="newborn"		
	1..146		
	147..2456		
	/note="acute phase response factor"		
	/codon_start=1		

[illegible]

QY	1728	CTGGCGGGGCTCTTGTGCAGCAATGAGTACGTACGTGAGAAAGACACACTGATGATGAAGAGCT	787
Db	806	GCTGGCGGGGCTCTTGTGCAGCAATGAGTACGTGAGAAAGACACACTGATGATGAAGAGCT	865
QY	788	GAGTCACTGGAAGAGGCGCGAGACATGCGCTGCATCGGAGGCGCCCTCCACACTGTGCGCT	847
Db	866	GGCTCACTGGAAGAGGCGCGAGACATGCGCGTGCATCGGAGGCGCCCTCCACACTGTGCGCT	925
QY	848	GGAGCGCTTGGAAAACTGATTACTTTCATTAGCAGAAATCTCAACTTCACACCCGCCAACA	907
Db	926	GGAGCGCTCTGGAAAACTGATTACTTTCATTAGCAGAAATCTCAACTTCAGACCCGCCAACA	985
QY	908	AATTAAAGAAACATGAGAGAGAGCTCAGACAGAAAGTGTCTCAAGAGGCGAGCCCTATCGTGA	967
Db	986	AATTAAAGAAACATGAGAGAGAGCTCAGACAGAAAGTGTCTCAAGAGGCGAGCCCTATCGTGA	1045
QY	968	GCACCGGCGCCATGCTGAGAGAGAGATGCTGGAGCGTTCAGAAACTTMAATGAAGATGC	1027
Db	1046	GCACCGGCGCCATGCTGAGAGAGAGATGCTGGAGCGTTCAGAAACTTMAATGAAGATGC	1105
QY	1028	CTTCTGTGTGAGAGCGGACCCCTGCATGCGCCATGCACCCGGACCGGCCCTTACTCATCAA	1087
Db	1106	CTTCTGTGTGAGAGCGGACCCCTGCATGCGCCATGCACCCGGACCGGCCCTTACTCATCAA	1165
QY	1088	GACTGGTGTCCAGTTTATCCACAGAAATCAGGTTGTGATGATTAATTCCTAGTTGAATTA	1147
Db	1166	GACTGGTGTCCAGTTTATCCACAGAAATCAGGTTGTGATGATTAATTCCTAGTTGAATTA	1225
QY	1148	TCAGCTTAAATTTAAAGTGTGCATTGCATATAAGACTCTGGGGATGTTGCTGCCCTCAGAGG	1207
Db	1226	TCAGCTTAAATTTAAAGTGTGTGCATTATAAAGACTCTGGGGATGTTGCTGCCCTCAGAGG	1285
QY	1208	GTCCTGGAATTTAACTATTCCTGGGCACGAACAAAAAGTATACATGAGAGAGTGTAA	1267
Db	1286	GTCCTGGAATTTAACTATTCCTGGGCACGAACAAAAAGTATATACATGAGAGAGTGTAA	1345
QY	1268	CAACGGCAGCTGTCTGTGCAGAGTTCAAGCACCTTGACCCCTTAGGGAGCAGAGATGTGGGAA	1327
Db	1346	CAACGGCAGCCTGTCTGTGCAGAGTTCAAGCACCTTGACCCCTTAGGGAGCAGAGATGTGGGAA	1405
QY	1328	TGGAGGCGCGTCCCAATTTGATGATCCCTCTTGATGCTACAGAGAGAGCTGCACCTGATCAC	1387
Db	1406	TGGAGGCGCGTCCCAATTTGATGATCCCTCTTGATGCTACAGAGAGCTGCACCTGATCAC	1465
QY	1388	CTTTCGAGACTGAGGTGTACACCAAGGCTCTCAAGATTGACCTTAGAGACCCACTCTTGGCC	1447
Db	1466	CTTTCGAGACTGAGGTGTACACCAAGGCTCTCAAGATTGACCTTAGAGACCCACTCTTGGCC	1525
QY	1448	AGTTGGTGTATCTCCAAACTCTGTGATGAGCCAAATGCTTGGGCATCAATCTGTGTGTA	1507
Db	1526	AGTTGGTGTATCTCCAAACTCTGTGATGAGCCAAATGCTTGGGCATCAATCTGTGTGTA	1585
QY	1508	TAAACATGCTGACCAATTAACCCCAAGAACGTGAACCTTTCACTAAGCGGCCAATTGGAAC	1567
Db	1586	TAAACATGCTGACCAATTAACCCCAAGAACGTGAACCTTTCACTAAGCGGCCAATTGGAAC	1645
QY	1568	CTGGAGACCAAGTGGCGAGGTGCTCAGCTGGAGATTCCTGTCACCAACCAAGCGAGGGCT	1627
Db	1646	CTGGAGACCAAGTGGCGAGGTGCTCAGCTGGAGATTCCTGTCACCAACCAAGCGGGGGCT	1705
QY	1628	GAGCATGAGACAGCTGACAAACGCTGGCTGAGAAAGCTCCTTAGGGCTGGTGTGAACCTACTC	1687
Db	1706	GAGCATGAGACAGCTGACAAACGCTGGCTGAGAAACCTCCTTAGGGCTGGTGTGAACCTACTC	1765
QY	1688	AGGGTGTCAATCATCAATGGGCTAAATTTCTCGAAAGAAAACTTGCTGGCAAGGGCTTCTC	1747
Db	1766	AGGGTGTCAATCATCAATGGGCTAAATTTCTCGAAAGAAAACTTGCTGGCAAGGGCTTCTC	1825
QY	1748	CTTCTGGGTGCTGACACAATATCATGCACTTGTGAAAAAGTATATCTTGGCCCTTTTG	1807
Db	1826	CTTCTGGGTGCTGCTAGACATATATCATGCACTTGTGAAAAAGTATATCTTGGCCCTTTTG	1885

QY	1808	GAATGAAGGATCATCATGGGTTTCATACGAAGAGGCGGGGCGCATCTCTAAAGAC	1887
Db	1886	GAATGAAGGATCACTCATGGGTTTCATACGAAGAGGCGGGGCGCATCTCTAAAGAC	1945
QY	1868	AAAGCCCCGGGACCTTCTACAGCGCTTCACGAGAGACACCAAAAGAGAGCGGCTAC	1927
Db	1946	AAAGCCCCGGGACCTTCTACTGCGCTTCACGAGAGACACCAAAAGAGAGGGGCTAC	2005
QY	1928	TTTTCACTTGGGTGAAAAAGACATCACTGGCAAGACCCAGATCCAGTCTGTAGCCCTA	1987
Db	2006	TTTTCACTTGGGTGAAAAAGACATCACTGGCAAGACCCAGATCCAGTCTGTAGCCCTA	2065
QY	1988	CACCAAGACAGACGTGAACAACATGCTATTTGCTGAATATCATGAGGCTATTAAGATAT	2047
Db	2066	CACCAAGACAGACGTGAACAACATGCTATTTGCTGAATATCATGAGGCTATTAAGATAT	2125
QY	2048	GGATCGGACCAACATCCGAGTGTCTCCACTTGTCTACCTCTACCCCGACATTTCCAAAGA	2107
Db	2126	GGATCGGACCAACATCCGAGTGTCTCCACTTGTCTACCTCTACCCCGACATTTCCAAAGA	2189
QY	2108	GGAGGCAATTTGGAAAGTACTGTAGGCCCGGAGCCGAGAGACACCCCGAAGCCGACCCAGG	2167
Db	2186	GGAGGCAATTTGGAAAGTACTGTAGGCCCGGAGCCGAGAGACACCCCGAAGCCGACCCAG	2244
QY	2168	TGAGTCGCGCCCGTACTCTGAAGACCAAGTTTCATCTGTGTACACCAAGACCTGACGAA	2227
Db	2245	- - GTGCGCCCGTACTCTGAAGACCAAGTTTCATCTGTGTACACCAAGACCTGACGAA	2302
QY	2228	TACCATTAACCTGCGGATGTGTCCCGCCGACCTTTAAGATCAATGATGACGTTTGGAAATTA	2287
Db	2303	TACCATTAACCTGCGGATGTGTCCCGCCGACCTTTAAGATCAATGATGACGTTTGGAAATTA	2367
QY	2288	CGGTGAAGGTGTGAAGCCCTAGCAGAGAGGGCAGTTTGAGTCCGCTCAGCTTTTGACATGA	2347
Db	2363	CGGTGAAGGTGTGTAGCCCTCAGAGAGAGGGCAGTTTGAGTCCGCTCAGCTTTTGACATGA	2422
QY	2348	TCTGACCTTGGAGTGTCTACTCTCCCATGTGTGAGAGCTGAAACCAAGAGCTGCAGAGA	2407
Db	2423	TCTGACCTTGGAGTGTCTACTCTCCCATGTGTGAGAGCTGAAACCAAGAGCTGCAGAGA	2487
QY	2408	CGGACATTGAGACCTGCCCCGAGCTACACCCCTAAGCAGCCGAAACCCCATATCGCTCG	2467
Db	2483	CGGACATTGAGACCTGCCCCGAGCTACACCCCTAAGCAGCCGAAACCCCATATCGCTCG	2547
QY	2468	AAACTCTCACTTTGTGGTCCAGA - TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT	2522
Db	2543	AAACTCTCACTTTGTGGTCCAGATTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT	2603
QY	2527	GGGCAATCTGGGCACTTTT TAAAGAGAGAAATGAGTGAAGTGGGAGTAAACTGTAT	2586
Db	2603	GGGCAATCTGGGCACTTTT TAAAGAGAGAAATGAGTGAAGTGGGAGTAAACTGTAT	2667
QY	2587	GTAAAGAGAGAG - ACCTCTGAGTCTGGGAGTGGGCTGAGACAGAGGAGGCAAGG	2641
Db	2663	GTAAAGAGAGAGACACTCTGAGTCTGGGAGTGGGCTGAGAGAGAGAGGAG - -CAAGG	2720
QY	2646	GGAACACCTCTCTGTCTGCGCGGCTGCGGCTCTCTTTTCAGACAGCTGGGGGTTGGTTT	2707
Db	2721	GGAACACCTCTCTGTCTGTCTGCGCGGCTCTCTTTTCACAGCTC - GGGGTTGGTTGTT	2777
QY	2706	AGACAGTGCCTCTGAGTGGCCCATGAGCTGTATCTGTGGCCCACTGTGAGCTGATACCC	2766
Db	2780	AGACAGTGCCTCTGTGTGTCAT - GCATCTGTGTGGCCCACTGTGTGAGCTGATACCC	2833
QY	2766	ATTCTGGGAACTCTGGCTGTGACATTTCAACTTGTCTAATATTCACATGAAGCTAGGA	2821
Db	2839	AGGCT - GGAACCTCTGGGCTGTGACTTTCAACTTGTCTAATATTCACATGAAGCTAGGA	2891
QY	2826	CTAAGCCGAGAGGTTCTCTTTAAATT 2853	
Db	2898	CTAAGCCCA - GAGGTTCTCTTTAAATT 2924	

RESULT	9				
LOCUS	188783				
DEFINITION	Sequence 7 from patent US 5719042.		PAT	10-AUG-1998	
ACCESSION	188783				
VERSION	188783.1	GI:3408723			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 2652)				
TITLE	Kishimoto,T. and Akira,S. Nucleic acids encoding transcription factor APRF (acute phase response factor) Patent: US 5719042-A 7 17-FEB-1998;				
JOURNAL	Location/Qualifiers				
FEATURES	1..2652				
source	/organism="unknown"				
BASE COUNT	683 a 706 c 748 g 515 t				
ORIGIN					

Query Match	84.2%	Score 2415.6	DB 6	Length 2652
Best Local Similarity	99.3%	Pred. No. 0		
Matches 2436: Conservative	0	Mismatches 14	Indels 2	Gaps 1
QY 13 CAGGCCGGCGCATCGGGC--TCAGCCCGGAGACAGTCGAGACCCTCGATCGCAGCAGAT	70			
Db 201 CAGGGCGGGCGCATCGGGGCTTCAGCCCGGAGACAGTCGAGACCCTCGACTCCAGCAGAT	260			
QY 71 GGCTCAGTGGAAACCAAGCTCAGCAGCTGGACACAGGCTACCTGGAAGCAGCTGCACAGCT	130			
Db 261 GGCTCAGTGGAAACCAAGCTCAGCAGCTGGACACAGGCTACCTGGAAGCAGCTGCACAGCT	320			
QY 131 GTACAGCGACACGTTCCCATGGAGCTGGGAGAGTTCTCGGACACCTTGATTTAGAGTCA	190			
Db 321 GTACAGCGACACGTTCCCATGGAGCTGGGAGAGTTCTCGGACACCTTGATTTAGAGTCA	380			
QY 191 ACACCTGGGATATGCAGCCAGCAAAAGTCACATGCCACGTTGGTGTTCATATCTCTT	250			
Db 381 ACACCTGGGATATGCAGCCAGCAAAAGTCACATGCCACGTTGGTGTTCATATCTCTT	440			
QY 251 GGGTGAATATTGACCAAGCAGATATAGCCGATCTCTGCAAGAGTCCAAATGTCCTATCAGCA	310			
Db 441 GGGTGAATATTGACCAAGCAGATATAGCCGATCTCTGCAAGAGTCCAAATGTCCTATCAGCA	500			
QY 311 CAACCTTGCAGAAATCAAGCACTTTCTGAGAGCAGATATCTTGAGAAACCATGGAAT	370			
Db 501 CAACCTTGCAGAAATCAAGCACTTTCTGAGAGCAGATATCTTGAGAAACCATGGAAT	560			
QY 371 TGCCTGGATCTGGCCCGATGCTGTGGGAAAGTCCGCTCCTCTCCACAGCGCAGCCAC	430			
Db 561 TGCCTGGATCTGGCCCGATGCTGTGGGAAAGTCCGCTCCTCTCCACAGCGCAGCCAC	620			
QY 431 GGCACCCCAAGAAAGGGGCGCAGGCCACACCCCAAGCGCCCGTATGTACAGAGAAGCA	490			
Db 621 GGCACCCCAAGAAAGGGGCGCAGGCCACACCCCAAGCGCTGCGTATGTACAGAGAAGCA	680			
QY 491 GCAGATGTTTGAGCAGCATCTTCAGAGATGTCGGAAGCGAGTGCAGAGATCTAGAACAGAA	550			
Db 681 GCAGATGTTTGAGCAGCATCTTCAGAGATGTCGGAAGCGAGTGCAGAGATCTAGAACAGAA	740			
QY 551 AATGAAGTGGTGGAGAAACCTCCAGGAGACGCTTGTATTTCAACTACAAAACCTTCAAGAG	610			
Db 741 AATGAAGTGGTGGAGAAACCTCCAGGAGACGCTTGTATTTCAACTACAAAACCTTCAAGAG	800			
QY 611 CCAAGGAGACATGCAGAGTCTGTAATGGAAGAACACAGTCTGTGACCACAGCAAGATGCA	670			
Db 801 CCAAGGAGACATGCAGAGTCTGTAATGGAAGAACACAGTCTGTGACCACAGCAAGATGCA	860			
QY 671 GCAGTGGAAACGATGCTCACAGCCCTGGACCCAGATCGGAGAAAGCATTTGTAGTGAGCT	730			
Db 861 GCAGTGGAAACGATGCTCACAGCCCTGGACCCAGATCGGAGAAAGCATTTGTAGTGAGCT	920			

QY	731	GGCGGGCGCTTTGTTCAGCAATATGGAGTACGTGCACAGAAAGACACTACTAGATGAAAGAGCTGGC	790
Db	921	GGCGGGCGCTTTGTTCAGCAATATGGAGTACGTGCACAGAAAGACACTACTAGATGAAAGAGCTGGC	980
QY	791	TGACTGAGAAAGCGCGCCAGAGATTCGGCTCATCGAGAGCGCTTCGCCAATATCGCTGGGA	850
Db	981	TGACTGAGAAAGCGCGCCAGAGATTCGGCTCATCGAGAGCGCTTCGCCAATATCGCTGGGA	1040
QY	851	CCGTCGTGAAAACTGGATTAACCTTCATTAGCAGAAATCTCAACTTCAGACCCGCCAACAAAT	910
Db	1041	CCGTCGTGAAAACTGGATTAACCTTCATTAGCAGAAATCTCAACTTCAGACCCGCCAACAAAT	1100
QY	911	TAAAGAACTGGAGAGAGCTGCCAGCAAGAAAGTCCCTACAAAGGGGAGCCATATGCTGAGCA	970
Db	1101	TAAAGAACTGGAGAGAGCTGCCAGCAAGAAAGTCCCTACAAAGGGGAGCCATATGCTGAGCA	1160
QY	971	CCGCGCCATTCGCTGGAGAGAGAGATTCGTGAGAGCTGTTCCACAACCTTAATGAAGAGTCCCT	1030
Db	1161	CCGCGCTTCGCTGGAGAGAGAGATTCGTGAGAGCTGTTCCACAACCTTAATGAAGAGTCCCT	1220
QY	1031	CGTGGTGCAGCGGCAGACCCCTTGCAATGCCCATGCAACCCGAGCCGCCCTTAAGTCATCAAGAC	1090
Db	1221	CGTGGTGCAGCGGCAGACCCCTTGCAATGCCCATGCAACCCGAGCCGCCCTTAAGTCATCAAGAC	1280
QY	1091	TGCTGTCCAGTTTACACAGAAAGTCAGAGTTGCTGGTAAATTTTCCAGATGGAATTATCA	1150
Db	1281	TGCTGTCCAGTTTACACAGAAAGTCAGAGTTGCTGGTAAATTTTCCAGATGGAATTATCA	1340
QY	1151	GCTTAAAAATTAAGTGTGCATTGATTAAGACCTCTGGGAGTGTTCGTCGCCCTCAGAGGCTC	1210
Db	1341	GCTTAAAAATTAAGTGTGCATTGATTAAGACCTCTGGGAGTGTTCGTCGCCCTCAGAGGCTC	1400
QY	1211	TGCGAAATTTAAACATTCCTGGGCGACAGAACAAAAAGTGAATGAACATGAGAGAGCTTAACAA	1270
Db	1401	TGCGAAATTTAAACATTCCTGGGCGACAGAACAAAAAGTGAATGAACATGAGAGAGCTTAACAA	1460
QY	1271	CGCGAGTCCTGCTCGAGAGTTCAAGCAACGACCTAGACCCTTAGGGAGCAGAGATGTGGGAATGG	1330
Db	1461	CGCGATCTGCTCGAGAGTTCAAGCAACGACCTAGACCCTTAGGGAGCAGAGATGTGGGAATGG	1520
QY	1331	AGGCGGTGCGCAATGTGTATGCTCTCCTTGATTCGTGACTGAGAGAGCTGCACCTATACACTT	1390
Db	1521	AGGCGGTGCGCAATGTGTATGCTCTCCTTGATTCGTGACTGAGAGAGCTGCACCTATACACTT	1580
QY	1391	CGAGACTAGGTTGACCAACCAAGCGCTCAAGATTGACGCTAGAACCCACACTCCCTGCCAGT	1450
Db	1581	CGAGACTAGGTTGACCAACCAAGCGCTCAAGATTGACGCTAGAACCCACACTCCCTGCCAGT	1640
QY	1451	TGTGTGATTCCTCAACATCTGTCCAGATGCCAAATGCTTGGGCAATCAATCCTGTGGTATTA	1510
Db	1641	TGTGTGATTCCTCAACATCTGTCCAGATGCCAAATGCTTGGGCAATCAATCCTGTGGTATTA	1700
QY	1511	CATGCTGACCAATTAACCCCAAGAACGTAAGCTTCTTCACATAACCGCCCAATTGGAAACCTG	1570
Db	1701	CATGCTGACCAATTAACCCCAAGAACGTAAGCTTCTTCACATAACCGCCCAATTGGAAACCTG	1760
QY	1571	GGACCAAGTGGCGAGAGTGTCTCAGCTGGCAGTTCTCGTCACCACCACCAAGCAGGAGCTGAG	1630
Db	1761	GGACCAAGTGGCGAGAGTGTCTCAGCTGGCAGTTCTCGTCACCACCACCAAGCAGGAGCTGAG	1820
QY	1631	CATGAGCAGCTGCAACAGCGCTGGCTGAGAAAGCTCTTAAGGCGCTGGCTGTAACTCTCAAGS	1690
Db	1821	CATGAGCAGCTGCAACAGCGCTGGCTGAGAAAGCTCTTAAGGCGCTGGCTGTAACTCTCAAGS	1880
QY	1691	GTGTGACATCAATGGGCTAAATTTCTGCAAAAGAAACATAGGCTGGCAGAGGCTTCTCCCT	1750
Db	1881	GTGTGACATCAATGGGCTAAATTTCTGCAAAAGAAACATAGGCTGGCAGAGGCTTCTCCCT	1940
QY	1751	CTGGGCTCTGGCTTAACAAATATCATGCACCTCTTGAAAAAAGTATATCTTGGCCCTTTGGAA	1810
Db	1941	CTGGGCTCTGGCTTAACAAATATCATGCACCTCTTGAAAAAAGTATATCTTGGCCCTTTGGAA	2000

Qy	1031	CGGGTGGACGGCAGCCCTTGCAATGCCCATGCAACCGGACCGGCCCTTAgTATCATCAAGC	1090
Db	1221	CGTGTGTGAGGGGAGGCCCTTGCAAGCCCATGCAACCGGACCGGCCCTTAgTATCATCAAGC	1280
Qy	1091	TGCTGTCCAGTTTACAGAGAAAGTCAGGTTGCGGTCAAAATTTCTGAGTGAATTATCA	1150
Db	1281	TGCTGTCCAGTTTACAGAGAAAGTCAGGTTGCGGTCAAAATTTCTGAGTGAATTATCA	1340
Qy	1151	GCTTAAAAATTAAAGTGTGCATTTGATTAAGACTCTGGGGATGTTGCTGCCCTCAGAGGTC	1210
Db	1341	GCTTAAAAATTAAAGTGTGCATTTGATTAAGACTCTGGGGATGTTGCTGCCCTCAGAGGTC	1400
Qy	1211	TGCGAAATTTAACTTCTGGGGACAGAACACAAAAGTGATTAACATGAGAGACTTAAACA	1270
Db	1401	TGCGAAATTTAACTTCTGGGGACAGAACACAAAAGTGATTAACATGAGAGACTTAAACA	1460
Qy	1271	CGGAGCTGTCTCGAGAGTTCAAGCACTGACCCTTAGGGAGCAGATGTGGGAATGG	1330
Db	1461	CGGAGCTGTCTCGAGAGTTCAAGCACTGACCCTTAGGGAGCAGATGTGGGAATGG	1520
Qy	1311	AGGCCGTGCCAATTGTATGTGCTCTTGATCTGTGACTGAGGAGCTGCACCTGATCACTT	1390
Db	1521	AGGCCGTGCCAATTGTATGTGCTCTTGATCTGTGACTGAGGAGCTGCACCTGATCACTT	1580
Qy	1391	CGAGACTAGGTTGACCAACCAAGGCCCTCAAGATTGAGCTAGAACCCACTCCTTGGCACT	1450
Db	1581	CGAGACTAGGTTGACCAACCAAGGCCCTCAAGATTGAGCTAGAACCCACTCCTTGGCACT	1640
Qy	1451	TGTGCTGATCTCCAAACATCTGTGCAGATGCCAAATGCTTGGGCTCAATCTCTGGATATA	1510
Db	1641	TGTGCTGATCTCCAAACATCTGTGCAGATGCCAAATGCTTGGGCTCAATCTCTGGATATA	1700
Qy	1511	CATGCTGACCAATTAACCCCAAGAACGTGAACCTTCTCACTAAGCCCGCAATTGGAACCTG	1570
Db	1701	CATGCTGACCAATTAACCCCAAGAACGTGAACCTTCTCACTAAGCCCGCAATTGGAACCTG	1760
Qy	1571	GGACCAACTGGCCGAGGTGCTCAAGCTGGGACGTTCTGTCTCACACCAACAGGGGGGCTAG	1630
Db	1761	GGACCAACTGGCCGAGGTGCTCAAGCTGGGACGTTCTGTCTCACACCAACAGGGGGGCTAG	1820
Qy	1631	CATCAGACAGCTGACAACGCTGGGTGAGAGCTCTAGAGGCCGTGGTGAATACTCTACGG	1690
Db	1821	CATCAGACAGCTGACAACGCTGGGTGAGAGCTCTAGAGGCCGTGGTGAATACTCTACGG	1880
Qy	1691	GTTGTACATCACTAGGCTAAATTTCTGCAAAAGAACATGGCTGGGCAAGGCTTCTCCTT	1750
Db	1881	GTTGTACATCACTAGGCTAAATTTCTGCAAAAGAACATGGCTGGGCAAGGCTTCTCCTT	1940
Qy	1751	CTGGGCTCTGGCTAACAATATCATTCACACCTTTTGAAAAAGTATATCTTGGCCCTTGGAA	1810
Db	1941	CTGGGCTCTGGCTAACAATATCATTCACACCTTTTGAAAAAGTATATCTTGGCCCTTGGAA	2000

QY 1811 TGAAGGTACATCATGGTTCATCAGCAGAGCGGGGACCTCTTAAGCAGAA 1870
| | | | |
Db 2001 TGAAGGTACATCATGGTTCATCAGCAGAGCGGGGACCTCTTAAGCAGAA 2060
QY 1871 GCGCCCGGACCTCTCTCACTGCGCTTCAGAGCAGCAGCAAGAAGAGGGGTCACCTT 1930
| | | | |
Db 2061 GCGCCCGGACCTCTCTCACTGCGCTTCAGAGCAGCAGCAAGAAGAGGGGTCACCTT 2120
QY 1931 CACTTGGGTGAAAAGAGCATCATGGGCAAGACCCAGATCCAGTCTGTAGAGCCATACAC 1990
| | | | |
Db 2121 CACTTGGGTGAAAAGAGCATCATGGGCAAGACCCAGATCCAGTCTGTAGAGCCATACAC 2180
QY 1991 CAAGCAGCAGCTGACCAACATGTCATTTGCTGAATCATATGGGGTATTAAGATCATGA 2050
| | | | |
Db 2181 CAAGCAGCAGCTGACCAACATGTCATTTGCTGAATCATATGGGGTATTAAGATCATGA 2240
QY 2051 TCGACCAACATCTGGTGTCTCCACTTGTCACTCTACCCGACATTTCCCAAGAGAGA 2110
| | | | |
Db 2241 TCGACCAACATCTGGTGTCTCCACTTGTCACTCTACCCGACATTTCCCAAGAGAGA 2300
QY 2111 GGCATTTGCAAACTACTGTAAGCGCCGAGAGCCAGACCCGAGCGGACCCAGCTAG 2170
| | | | |
Db 2301 GGCATTTGCAAACTACTGTAAGCGCCGAGAGCCAGACCCGAGCGGACCCAGCTAG 2360
QY 2171 TCGTCCCGGTACTGTAAGCAAGTTCACTGTGTGACACCAACGACCTGACAGCAATAC 2230
| | | | |
Db 2361 TCGTCCCGGTACTGTAAGCAAGTTCACTGTGTGACACCAACGACCTGACAGCAATAC 2420
QY 2231 CATTGACCTGCGATGTCCCCCGGACATTTAGATTGATGATGACATTTTGAATTAACCG 2290
| | | | |
Db 2421 CATTGACCTGCGATGTCCCCCGGACATTTAGATTGATGATGACATTTTGAATTAACCG 2480
QY 2291 TGAAGGTGCGAGCCCGCAGCAGAGGCGAGTTGAGTCCCTCAGCTTGACATGATCT 2350
| | | | |
Db 2481 TGAAGGTGCGAGCCCGCAGCAGAGGCGAGTTGAGTCCCTCAGCTTGACATGATCT 2540
QY 2351 GACCTGCGAGTGTCTACCTCTCCCATGTGAGAGCTGAACCAAGAGCTGACAGACGT 2410
| | | | |
Db 2541 GACCTGCGAGTGTCTACCTCTCCCATGTGAGAGCTGAACCAAGAGCTGACAGACGT 2600
QY 2411 GACTTGAGAACCTGCGCGCTGTCCAGCCCTTAAGCAGCGCAACCCCATATC 2462
| | | | |
Db 2601 GACTTGAGAACCTGCGCGCTGTCCAGCCCTTAAGCAGCGCAACCCCATATC 2652
RESULT 10
LOCUS 188784 2652 bp DNA PAT 10-AUG-1998
DEFINITION Sequence 8 from patent US 5719042.
ACCESSION 188784
VERSION 188784.1 GI:3408724
KEYWORDS
SOURCE .
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2652)
AUTHORS Kishimoto,T. and Akira,S.
TITLE Nucleic acids encoding transcription factor APRF (acute phase response factor)
JOURNAL Patent: US 5719042-A 8 17-FEB-1998;
FEATURES
source 1..2652
BASE COUNT 683 a 706 c 748 g 515 t
ORIGIN
Query Match 84.2% Score 2415.6; DB 6; Length 2652;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2436; Conservative 0; Mismatches 14; Indels 2; Gaps 1;
QY 13 CAGCGCGGCGAGTGGGCG -TCAGCCCGGAGACAGTGGAGACCCCTGACTGCACAGAGAT 70
| | | | |

Db 201 CAGCGCGGCGAGTGGGCGCTCAGCCCGAGACAGTGGAGACCCCTGACTGCACAGAGAT 260
QY 71 GGCTAGTGGAAACAGCTGTCACACCTGAGACACAGCTACCTGAAAGAGAGTGCACACT 130
| | | | |
Db 261 GGCTAGTGGAAACAGCTGTCACACCTGAGACACAGCTACCTGAAAGAGAGTGCACACT 320
QY 131 GTACAGCAGACAGTTCCTCCATGAGCTGCGGAGCTGTCTCTGACCTTGGATTGAGATCA 190
| | | | |
Db 321 GTACAGCAGACAGTTCCTCCATGAGCTGCGGAGCTGTCTCTGACCTTGGATTGAGATCA 380
QY 191 AGACTGGGCATATGACAGCCAGCAAGAGATGCATGTCACAGCTTGGTGTTCATTAATCTCT 250
| | | | |
Db 381 AGACTGGGCATATGACAGCCAGCAAGAGATGCATGTCACAGCTTGGTGTTCATTAATCTCT 440
QY 251 GGGTGAATTTGACCAAGATATAGCGATTCTCTGCAAGAGTCCAAATGTCCTTATCACA 310
| | | | |
Db 441 GGCTC AATTGACCAAGATATAGCGATTCTCTGCAAGAGTCCAAATGTCCTTATCACA 500
QY 311 CAACCTTGCAAGATTCAGACACTTCTGCAAGAGAGATATCTTGAGAAAGCCAAATGGAAT 370
| | | | |
Db 501 CAACCTTGCAAGATTCAGACACTTCTGCAAGAGAGATATCTTGAGAAAGCCAAATGGAAT 560
QY 371 TGGCCGAGTCTGGGCGGATGCTGTGGGAAAGTCTGCTCTCCAGACGCGAGCCAC 430
| | | | |
Db 561 TGGCCGAGTCTGGGCGGATGCTGTGGGAAAGTCTGCTCTCCAGACGCGAGCCAC 620
QY 431 GGCAGCCGCAAGAGGCGGAGCCAGCCCAACCCAGCCGCGGTAGTGCAGAGAGCA 490
| | | | |
Db 621 GGCAGCCGCAAGAGGCGGAGCCAGCCCAACCCAGCCGCGGTAGTGCAGAGAGCA 680
QY 491 GCAGATGTTGGAGCAGCATCTTTCAGAGATCCGGAAGCCAGTGCAGATCTGAAGACAGA 550
| | | | |
Db 681 GCAGATGTTGGAGCAGCATCTTTCAGAGATCCGGAAGCCAGTGCAGAGTCTGAAGACAGA 740
QY 551 AATGAAGTGTGGAGAACCTCCAGAGCAGATTTGATTTCACTTACAAACCCCTCAAGAG 610
| | | | |
Db 741 AATGAAGTGTGGAGAACCTCCAGAGCAGATTTGATTTCACTTACAAACCCCTCAAGAG 800
QY 611 CCAAGAGACATGACAGATCTGAATGAAACCAACAGCTCTGTACCAAGACAGAAATGCA 670
| | | | |
Db 801 CCAAGAGACATGACAGATCTGAATGAAACCAACAGCTCTGTGTACCAAGACAGAAATGCA 860
QY 671 GCAGCTGGAAACAGATGCTCAGACCTCGAGCCAGATGCGGAGAGAGCATGTGAGTGA 730
| | | | |
Db 861 GCAGCTGGAAACAGATGCTCAGACCTCGAGCCAGATGCGGAGAGAGCATGTGAGTGA 920
QY 731 GCGGGGCTCTTGTCAAGCAATGGAGTACGTGCAAGAGACACTGATGATGAGAGCTGGC 790
| | | | |
Db 921 GCGGGGCTCTTGTCAAGCAATGGAGTACGTGCAAGAGACACTGATGATGAGAGCTGGC 980
QY 791 TGACTGGAAGAGCGCGCCAGAGATCGGTGTCATCGAGAGCGCCGCCAACAATCGCCTGCA 850
| | | | |
Db 981 TGACTGGAAGAGCGCGCCAGAGATCGGTGTCATCGAGAGCGCCGCCAACAATCGCCTGCA 1040
QY 851 CCGTCTGGAAGACTGATTAATCTTATAGCAGATCTCAACTTCAGACCCGCCAACAAT 910
| | | | |
Db 1041 CCGTCTGGAAGACTGATTAATCTTATAGCAGATCTCAACTTCAGACCCGCCAACAAT 1100
QY 911 TAAGAACTGGAGAGCTGCAAGCAAGAAAGTCTCTACAAAGGGGAGCTTATGTCAGAGA 970
| | | | |
Db 1101 TAAGAACTGGAGAGCTGCAAGCAAGAAAGTCTCTACAAAGGGGAGCTTATGTCAGAGA 1160
QY 971 CCGGCCATGCTGGAGAGAGAGATCGTGAGAGCTGTGAGAACTTAATGAAGAGTCCCTT 1030
| | | | |
Db 1161 CCGGCCATGCTGGAGAGAGAGATCGTGAGAGCTGTGAGAACTTAATGAAGAGTCCCTT 1220
QY 1031 CCGTGTGAGAGCGGAGGCTTCATGCCATGCAACCGGAGCCGCTTACTATCAAGAC 1090
| | | | |
Db 1221 CCGTGTGAGAGCGGAGGCTTCATGCCATGCAACCGGAGCCGCTTACTATCAAGAC 1280
QY 1091 TCGTGTCACTTTACCAAGCAAGTCAAGTGTGCTGCAAAATTCAGTGTGAATATCA 1150
| | | | |
Db 1281 TCGTGTCACTTTACCAAGCAAGTCAAGTGTGCTGCAAAATTCAGTGTGAATATCA 1340
| | | | |

1151 GCTTAAATTAAGTGTGATGATTAAGACTGTGGGATGTTGCTGCCCCAGAGGCT 1210
1341 GCTTAAATTAAGTGTGATGATTAAGACTGTGGGATGTTGCTGCCCCAGAGGCT 1400
1211 TCGGAATTTAATCTTGTGGGACGACAAAGTGAATGAAACATGAGAGAGTCTAACAA 1270
1401 TCGGAATTTAATCTTGTGGGACGACAAAGTGAATGAAACATGAGAGAGTCTAACAA 1460
1271 GCGAGGCTGCTGTGAGAGTGAAGACCTGACCCCTTAGGAGACAGATGTTGGGAATG 1330
1461 GCGAGGCTGCTGTGAGAGTGAAGACCTGACCCCTTAGGAGACAGATGTTGGGAATG 1520
1331 AGGCGGTGCAATTTGATGATGCTCTTGTATGCTGACTGAGAGCTGACCTGATCCTT 1390
1521 AGGCGGTGCAATTTGATGATGCTCTTGTATGCTGACTGAGAGCTGACCTGATCCTT 1580
1391 CGAGACTGAGGTGATGACCAAGGCTCAAGATGACCTAGAGACCTGACCTGCTGCTG 1450
1581 CGAGACTGAGGTGATGACCAAGGCTCAAGATGACCTAGAGACCTGACCTGCTGCTG 1640
1451 TGTGATGATCTCCAAACATCTGTGATGCTGCAATGCTTGGGATCAATCTGTGTATAA 1510
1641 TGTGATGATCTCCAAACATCTGTGATGCTGCAATGCTTGGGATCAATCTGTGTATAA 1700
1511 CATGCTGACCAATTAAGCCCAAGAGAGCTGAACTCTCAATAGCGGCAATGGAACCTG 1570
1701 CATGCTGACCAATTAAGCCCAAGAGAGCTGAACTCTCAATAGCGGCAATGGAACCTG 1760
1571 GAGCAAGTGGCGAGAGTGTCTGAGTGGAGTTCCTGCTCCACCAAGAGAGGAGCTGAG 1630
1761 GAGCAAGTGGCGAGAGTGTCTGAGTGGAGTTCCTGCTCCACCAAGAGAGGAGCTGAG 1820
1631 CATGAGAGAGTGTGACAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1690
1821 CATGAGAGAGTGTGACAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1880
1691 GTGTGATGATGATGAGGCTTAAATTTGCAAGAAACATGCTGAGAGGCTTCTCCTT 1750
1881 GTGTGATGATGATGAGGCTTAAATTTGCAAGAAACATGCTGAGAGGCTTCTCCTT 1940
1751 CTGGGCTGTGCTGACAAATATCATGACCTTGTGAAAAGTATATCTTGCCCTTTGGA 1810
1941 CTGGGCTGTGCTGACAAATATCATGACCTTGTGAAAAGTATATCTTGCCCTTTGGA 2000
1811 TGAAGGTATCATGATGAGTGTTCATGACAAAGAGAGGAGGAGGAGGAGGAGGAGG 1870
2001 TGAAGGTATCATGATGAGTGTTCATGACAAAGAGAGGAGGAGGAGGAGGAGGAGG 2060
1871 GGGCGGGGACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1930
2061 GGGCGGGGACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2120
1931 CACTTGGTGAAG 1990
2121 CACTTGGTGAAG 2180
1991 CAAG 2050
2181 CAAG 2240
2051 TGGGAG 2110
2241 TGGGAG 2300
2111 GGCATTTGAAAAGTACTGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2170
2301 GGCATTTGAAAAGTACTGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2360
2171 TGTGCCCCGTACCTGAG 2230
2361 TGTGCCCCGTACCTGAG 2420

2231 CATTGACCTGCCGATGTCCTCCCGACCTTTAGATTCATTGATGACGTTGGAAATACG 2290
2421 CATTGACCTGCCGATGTCCTCCCGACCTTTAGATTCATTGATGACGTTGGAAATACG 2480
2291 TGAAGTGTGAG 2350
2481 TGAAGTGTGAG 2540
2351 GACCTGAG 2410
2541 GACCTGAG 2600
2411 GACTTGAAG 2462
2601 GACTTGAAG 2652

RESULT 11

MUSAPRF

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

FEATURES

source

1. 2652

Location/Qualifiers

/organism="Mus musculus"

/strain="BALB/c"

/sub_species="domesticus"

/db_xref="taxon:10090"

/issue_type="liver"

/issue_11b="lambda gll1 cdna library"

259..2571

/gene="APRF"

259..2571

/gene="APRF"

/note="ISGF3 p91-related transcription factor"

/function="transcription factor involved in the

gp130-mediated signalling pathway"

/product="acute phase response factor"

/protein_id="AA037254.1"

/db_xref="GI:476716"

/translation="MAQNMNLOLDITRYLEQLHQLYSDSPFMEQLRPLAPMIESODMA

VAARESHATLVEFNLLEIDDOYSRFLQESNVLYOHNLRIRKQFLQSRVLEKPEIA

RIVARESHATLVEFNLLEIDDOYSRFLQESNVLYOHNLRIRKQFLQSRVLEKPEIA

KMKVVENLDDDFPNYKTLKSGDDMDKNNQSVYRQKMOOLEOMLALDOMRSIV

SELACILSMEVYVOKTIDDEPILDMKRRQIACIGGPNICLDRENNITSLASOLG

TRQOIKRLELDQKVSFKDPIVORHPLLEEVILFRLMKSAFVYVRQCMHPD

RPLVLTQVQFTTKVRLVLEPELVQIKIVCIDKDSGDVAALGSKFNLVLTNRK

VINMESNNGSLAEFKHLLTREQRCNGGRANDASLIVEELHLITFEVEVHGL

KIDLETSLPVLVVISNICMPNMAASILMYNMLTNPNKVNVEFTKPIGTDQVAEVL

SMOESSTKRGSLIDQLTAEKLLGPGVNSGCOITMAKCKEKMAKGSFVWMLD

NIIIDYKIIYLAINEGYIMGFSKREERAILSTKPGPIILRESSEKSEGVETVY

EKDISKTQISQVETTKQOOLNNMSPAEIIMKTIKMDATNIIYSYLYVLDIYDIEEA

FKYICRPESQEHPEADPGSAAPLYLTKFLICVPTTCSNTIDLPMSPTLDSLMOPGN

GGAEPSSAGQESLTFDDMTSEATSPM"

BASE COUNT

ORIGIN

683 a 706 c 748 g 515 t

Query Match 84.2% Score 2415.6 DB 10: Length 2652:
Best Local Similarity 99.3% Pred. No. 0:
Matches 2436; Conservative 0; Mismatches 14; Indels 2; Gaps 1:

```

Oy 13 CAGGCGGCGCGAGTGGGGG--TCAGCGCGGAGACAGTGCAGAGCCCTGACTCAGTAGAGAT 70
Db 201 CAGGCGGCGCGAGTGGGGGCTTACGCGCGCGAGAGAGTCCAGAGAGCCCTGACTCAGTAGAGAT 260
Oy 71 GGCCTAGTGAACACAGCTTCGACAGCAGCTTGACACAGCTTACCTGAAAGCAGCTGCACACCT 130
Db 261 GGCCTAGTGAACACAGCTTGACAGCAGCTTGACACAGCTTACCTGAAAGCAGCTGCACACCT 320
Oy 131 GTACAGCCACAGCTTCCCATGAGAGCTGCGGAGTTCCTGCGAGCTTGAGATTGAGAGTCA 190
Db 321 GTACAGCCACAGCTTCCCATGAGAGCTGCGGAGTTCCTGCGAGCTTGAGATTGAGAGTCA 380
Oy 191 AGACTGGCATATGACAGCCAGCAAGAGTACATGCGACAGCTTGTTGTTGATTAATCTCTT 250
Db 381 AGACTGGCATATGACAGCCAGCAAGAGTACATGCGACAGCTTGTTGTTGATTAATCTCTT 440
Oy 251 GGGTGAATTTGACCAAGCAATATAGCCGATTCTGCAAGAGTCCATATGCTTATCAGCA 310
Db 441 GGGTGAATTTGACCAAGCAATATAGCCGATTCTGCAAGAGTCCATATGCTTATCAGCA 500
Oy 311 CAACCTTGAAGATACAGCAGTTCCTGAGAGCAGTATCTTGAGAGCAACCAATGGAAT 370
Db 501 CAACCTTGAAGATACAGCAGTTCCTGAGAGCAGTATCTTGAGAGCAACCAATGGAAT 560
Oy 371 TGCCCGGATCGTGGCCGATGCTGTGGAGAGTCTGCGCTCTCCAGAGCGGACGCCAC 430
Db 561 TGCCCGGATCGTGGCCGATGCTGTGGAGAGTCTGCGCTCTCCAGAGCGGACGCCAC 620
Oy 431 GCGAGCCCGCAAGGGGGGCGAGGCCAACCACCCAGCGCGCGTGTGAGAGAGCA 490
Db 621 GCGAGCCCGCAAGGGGGGCGAGGCCAACCACCCAGCGCGCGTGTGAGAGAGCA 680
Oy 491 GCAGATGTTGAGACAGCATCTTTCAGAGTGTCCGAAACGAGTGCAGATCTAGACAGAA 550
Db 681 GCAGATGTTGAGACAGCATCTTTCAGAGTGTCCGAAACGAGTGCAGATCTAGACAGAA 740
Oy 551 AATGAAGTGTGAGACAGCATCTTTCAGAGTGTCCGAAACGAGTGCAGATCTAGACAGAA 610
Db 741 AATGAAGTGTGAGACAGCATCTTTCAGAGTGTCCGAAACGAGTGCAGATCTAGACAGAA 800
Oy 611 CCAAGGAGACATGAGATCTGAAATGAAACCAACAGCTCTGTGACAGACAGAGATGCA 670
Db 801 CCAAGGAGACATGAGATCTGAAATGAAACCAACAGCTCTGTGACAGACAGAGATGCA 860
Oy 671 GCAGCTGGAACAGATGCTCAGACCCCTGGACCGAGATGCGGAGAAAGCATTTGTAGTGCT 730
Db 861 GCAGCTGGAACAGATGCTCAGACCCCTGGACCGAGATGCGGAGAAAGCATTTGTAGTGCT 920
Oy 731 GCGGGGCTCTTGTACAGCATGAGTACGTGACAGAAAGACATGACTGATGAAGAGCTGGC 790
Db 921 GCGGGGCTCTTGTACAGCATGAGTACGTGACAGAAAGACATGACTGATGAAGAGCTGGC 980
Oy 791 TGACTGAGAGGCGGCCAGAGATCGCTGATGAGAGGCGCTTCCCAACATCTGCTTGA 850
Db 981 TGACTGAGAGGCGGCCAGAGATCGCTGATGAGAGGCGCTTCCCAACATCTGCTTGA 1040
Oy 851 CCGTCTGGAAGAACTGATTAATCTTATGAGATCTCAACTTACAGCCCGCCCAACAAAT 910
Db 1041 CCGTCTGGAAGAACTGATTAATCTTATGAGATCTCAACTTACAGCCCGCCCAACAAAT 1100
Oy 911 TAAGAACTGAGAGAGCTGCAGCAAGAGTGTCTTACAGGGGCAACCTATCTGCAGCA 970
Db 1101 TAAGAACTGAGAGAGCTGCAGCAAGAGTGTCTTACAGGGGCAACCTATCTGCAGCA 1160
Oy 971 CCGGCGCATGCTGAGAGAGAGATCGTGGAGCTGTTCAGAAACTTAAATGAAGATGCTT 1030
Db 1161 CCGGCGCATGCTGAGAGAGAGATCGTGGAGCTGTTCAGAAACTTAAATGAAGATGCTT 1220
```

```

Oy 1031 CGTGTGAGGCGGAGGCGCTGACATGCCATGACACCGGAGCGGCCCTTACTCATCAAGAC 1090
Db 1221 CGTGTGAGGCGGAGGCGCTGACATGCCATGACACCGGAGCGGCCCTTACTCATCAAGAC 1280
Oy 1091 TGTGTCCAGTTTACACAGAAAGTACAGTGTGCTGTCAAATTTCTGAGTGAATTTACA 1150
Db 1281 TGTGTCCAGTTTACACAGAAAGTACAGTGTGCTGTCAAATTTCTGAGTGAATTTACA 1340
Oy 1151 GCTTAAATTTAAAGTGTGCTAATGATTAAGACTCTGGGAGTGTGCTGCCCTCAGAGGCTC 1210
Db 1341 GCTTAAATTTAAAGTGTGCTAATGATTAAGACTCTGGGAGTGTGCTGCCCTCAGAGGCTC 1400
Oy 1211 TCGAAATTTTAACTTGTGGGACAGACACAAAAGTATGAAACATGAGAGGATCTTAACAA 1270
Db 1401 TCGAAATTTTAACTTGTGGGACAGACACAAAAGTATGAAACATGAGAGGATCTTAACAA 1460
Oy 1271 GCGGAGCGCTGTGAGAGTTTCAACACCTTACACCTTATAGGAGAGAGATGTGGAAATGG 1330
Db 1461 GCGGAGCGCTGTGAGAGTTTCAACACCTTACACCTTATAGGAGAGAGATGTGGAAATGG 1520
Oy 1331 AGGCGGTGCAATTTGATGCTCTCTGATCGTGAAGTGAAGAGCTGCACCTGATACCTT 1390
Db 1521 AGGCGGTGCAATTTGATGCTCTCTGATCGTGAAGTGAAGAGCTGCACCTGATACCTT 1580
Oy 1391 CGAGACTGAGTGTACCAACCAAGGCGCTCAAGATGACCTAGAGAGCCACCTTCTGCCAOT 1450
Db 1581 CGAGACTGAGTGTACCAACCAAGGCGCTCAAGATGACCTAGAGAGCCACCTTCTGCCAOT 1640
Oy 1451 TGTGTGATCTCAACATCTGTCAAGATGCAAGATGCTTGGGAGTCAATCTGTGATTA 1510
Db 1641 TGTGTGATCTCAACATCTGTCAAGATGCAAGATGCTTGGGAGTCAATCTGTGATTA 1700
Oy 1511 CATGCTGACCAATTAACCCCAAGACGCTGAACCTTCTTCAATAGCCCGCAATTGGAACCTG 1570
Db 1701 CATGCTGACCAATTAACCCCAAGACGCTGAACCTTCTTCAATAGCCCGCAATTGGAACCTG 1760
Oy 1571 GGCACCAAGTGGCGGAGGCTGCTGAGTGTCTGCTCACACCAACGAGAGGCTGAG 1630
Db 1761 GGCACCAAGTGGCGGAGGCTGCTGAGTGTCTGCTCACACCAACGAGAGGCTGAG 1820
Oy 1631 CATGAGACAGCTGACAAAGCTGCTGAGAAAGCTTCTAGGGCTTGTGATCTACCTCAGG 1690
Db 1821 CATGAGACAGCTGACAAAGCTGCTGAGAAAGCTTCTAGGGCTTGTGATCTACCTCAGG 1880
Oy 1691 GTGTGATGATCAATGGGCTAAATTTGCAAAAGAAACATGGCTGGCAAGGCTTCTCTT 1750
Db 1881 GTGTGATGATCAATGGGCTAAATTTGCAAAAGAAACATGGCTGGCAAGGCTTCTCTT 1940
Oy 1751 CTGGGCTGTGCTAGACAAATATCATGACCTTGTGAAGAAAGTATATCTTGGCCCTTTGGA 1810
Db 1941 CTGGGCTGTGCTAGACAAATATCATGACCTTGTGAAGAAAGTATATCTTGGCCCTTTGGA 2000
Oy 1811 TGAAGGTACATATGCTTTCATGAGCAAGAGCGGAGCGGAGCCATCTTAAGCAACAA 1870
Db 2001 TGAAGGTACATATGCTTTCATGAGCAAGAGCGGAGCGGAGCCATCTTAAGCAACAA 2060
Oy 1871 GCCCGCGGACCTTCTTACTTGCCTTCAAGGAGAGAGCAAGAAAGAGAGGCTTCACTTT 1930
Db 2061 GCCCGCGGACCTTCTTACTTGCCTTCAAGGAGAGAGCAAGAAAGAGAGGCTTCACTTT 2120
Oy 1931 CACTTGGGTGAAAAGACATTCAGTGTGCAAGACCCAGATCTCTTAAAGCAATPAC 1990
Db 2121 CACTTGGGTGAAAAGACATTCAGTGTGCAAGACCCAGATCTCTTAAAGCAATPAC 2180
Oy 1991 CAAGCAGCAGCTGAAACATATGCTATTTGTAAGATCATGAGGCTATAGATCATGGA 2050
Db 2181 CAAGCAGCAGCTGAAACATATGCTATTTGTAAGATCATGAGGCTATAGATCATGGA 2240
Oy 2051 TGCAGACCAATCTGTGTCTCACTTGTCTACCTTACCCGAGCAATTTCCCAAGAGGA 2110
Db 2241 TGCAGACCAATCTGTGTGTCTCACTTGTCTACCTTACCCGAGCAATTTCCCAAGAGGA 2300
```

Qy	2111	GGATTTTGGAAAGACTGTATGGCCCGAGAGACCAAGACACCCCGGAAGCCAGCAACCGAGTAG	2170
Db	2301	GGCAATTTGGAAAGACTGTATGSCCCGAGAGACCAAGAGACACCCCGAAGCCGACCGAGTAG	2360
Qy	2171	TGCTGCCCCGACTACCTGAAGACCAAGTTTCATCTGTGTGACACCAACGACCTCGACCAATAC	2230
Db	2361	TGCTGCCCCGACTACCTGAAGACCAAGTTTCATCTGTGTGACACCAACGACCTCGACCAATAC	2420
Qy	2231	CATTGACCTGCCCATGTGCCCCCGGCACTTAAATTCATTGATGACAGTTTGGAAATTAACGG	2290
Db	2421	CATTGACCTGCCCATGTGCCCCCGGCACTTAAATTCATTGATGACAGTTTGGAAATTAACGG	2480
Qy	2291	TGAAGTGCTGAGCCCTCAGACAGAGAGGCAAGTTTAGTGCGTCACAGTTTGGACATGATCT	2350
Db	2481	TGAAGTGCTGAGCCCTCAGACAGAGAGGCAAGTTTAGTGCGTCACAGTTTGGACATGATCT	2540
Qy	2351	GACCTGGAGTGCTACGTCGCCCCCGCAATGTGAGAGAGCTGAAACACAGACCTCGACAGAGCT	2410
Db	2541	GACCTGGAGTGCTACGTCGCCCCCGCAATGTGAGAGAGCTGAAACACAGACCTCGACAGAGCT	2600
Qy	2411	GACTTGAACACACTGCCCCGCTGCTCCACCCCTTAAGCAGCCGGAACCCCATATC	2462
Db	2601	GACTTGAACACACTGCCCCGCTGCTCCACCCCTTAAGCAGCCGGAACCCCATATC	2652

Query Match	Best Local Similarity	84.2%	Score 2415.6;	DB 22;	length 2652;
Matches 2436;	Conservative	99.3%;	Pred. No. 0;	Mismatches 14;	Indels 2; Gaps 1;
QY	13 CAGGGCGGCCATTCGGG--TCAGCCGGAGACATCGAGACCCCTGCATCGACGAGAT				70
DB	201 CAGGGCGGCGAATCGGGGCTCAGCCCGGAGACATCGAGACCCCTGACTGACGAGCAT				260
QY	71 GGCTAGTGGACCAACGCTGCACACCACTGTGGACACAGCTTAACGTAAGCAAGCTACCACT				130
DB	261 GGCTAGTGGACCAACGCTGCACACCACTGTGGACACAGCTTAACGTAAGCAAGCTACCACT				320
QY	131 GTACAGCGACACGTTCCCATATGAGACTCTGGCAGTTCTTGGCACCTTGATTGAGAGTCA				190
DB	321 GTACAGCGACACCTTCCCATATGAGACTCTGGCAGTTCTTGGCACCTTGATTGAGAGTCA				380

RESULT	12		
EI0408			
ID	EI0408	standard; RNA; ROD; 2652 BP.	
XX			
AC	EI0408;		
XX			
SV	EI0408.1		
XX			
DT	08-OCT-1997 (Rel. 52, Created)		
DT	02-SEP-2000 (Rel. 65, Last updated, Version 2)		
XX			
DE	cDNA encoding an acute phase response factor.		
XX			
KW	JP 1995324096-A/2.		
XX			
OS	Mus sp.		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;		
OC	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
XX			
RN	[1]		
RP	1-2652		
RA	Shinriyou S., Kishimoto C.;		
RT	"TRANSCRIPT APRF";		
RL	Patent number JP1995324096-A/2, 12-DEC-1995.		
RL	KISHIMOTO CHUZO.		
XX			
OS	Mus sp. (mouse)		
CC	PN JP 1995324096-A/2		
CC	PD 12-DEC-1995		
CC	PF 03-APR-1995 JP 1995077584		
CC	PR 04-APR-1994 JP 94P 65825		
CC	PI SHINRIYOU SHIZUO, KISHIMOTO CHUZO		
CC	C07K14/52,A61K31/70,A61K38/00,A61K39/395,A61K48/00,C07H21/04,		
CC	C12N5/09,		
CC	PC C12P21/02/C12M1/21,C12N5/10,(C12P21/02,C12R1:91),(C12M1/21,		
CC	PC C12P21/02/C12M1/21,C12N5/10,C12R1:91),		
CC	PC C12N5/10,C12R1:91);		
CC	strandedness: Double;		
CC	topology: Linear;		
CC	key	Location/Qualifiers	
CC	FH		
CC	FH		
CC	FT	source	
CC	FT	1..2652	
CC	FT	/organism="Mus sp. "	
CC	FT	/tissue_type="liver"	
CC	FT	259..2571	
CC	FT	/product="APRF"	
XX			
Key		Location/Qualifiers	

QY	251	GGGTAAATTTGACCGACGATATATAGCCGATTTCTCGAAGAGTCCAAATGTCTCTATCAGCA	310
Db	441	GGGTAAATTTGACCGACGATATATAGCCGATTTCTCGAAGAGTCCAAATGTCTCTATCAGCA	500
QY	311	CAACCTTGCAGAGATCAAGCAGTTTCTCGAGAGCAGTATCTTTGAGAACCAATGGAAT	370
Db	501	CAACCTTGCAGAGATCAAGCAGTTTCTCGAGAGCAGTATCTTTGAGAACCAATGGAAT	560
QY	371	TGCCCCGATTCGTGGCCCCATGCTGTGGGAAGATCTGGCCCTCCGACGGCGCCAC	430
Db	561	TGCCCGCATCTGGCCCCATGCTGTGGGAAGATCTGGCCCTCCGACGGCGCCAC	620
QY	431	GGCAGCCCCGCAAGGGGGCCAGGGCCACACCACCAACAGCCGCCCTGTAGTGACAGAGACCA	490
Db	621	GGCAGCCCCAGAGGGGGGGCAGGGCCACACCACCAACAGCTGCCGTATGTAGACAGAGAACCA	680
QY	491	GCAGATGTTTGGACGACGACATCTTCAGAGATGTCCGGAAGCCAGTGACAGATCTTAGAACAGAA	550
Db	681	GCAGATGTTTGGACGACGACATCTTCAGAGATGTCCGGAAGCCAGTGACAGATCTTAGAACAGAA	740
QY	551	AATGAAAGGTGGTGGGAAGACCTCCAGAGAGACTTGTATTCAACTCAAAACCCCTCAAGAG	610
Db	741	AATGAAAGGTGGTGGGAAGACCTCCAGAGAGACTTGTATTCAACTCAAAACCCCTCAAGAG	800
QY	611	CCAGAGAGCATGACGAGATCTGTCAATGGGAACCAACCACTGTGTGACACAGAGAGATGCA	670
Db	801	CCAGAGAGCATGACGAGATCTGTCAATGGGAACCAACCACTGTGTGACACAGAGAGATGCA	860
QY	671	GCAGCTGCAACGATGCTTCACAGCCCTTGACCAGATGCGGAGAGCAATTTGTAGTAGCT	730
Db	861	GCAGCTGCAACGATGCTTCACAGCCCTTGACCAGATGCGGAGAGCAATTTGTAGTAGACT	920
QY	731	GGCGGGGCTTGTTCAGCAATGTAGTACGTCCAGAAAGCACTGACTGATGGAAGAGCTGGC	790
Db	921	GGCGGGGCTTGTTCAGCAATGTAGTACGTCCAGAAAGCACTGACTGATGGAAGAGCTGGC	980
QY	791	TGACGTGGAAGAGGGCGCCAGAGATGCGCTGTGATCGAGAGCCCTCCCAACATCTCCCTGGA	850
Db	981	TGACGTGGAAGAGGGCGCCAGAGATGCGCTGTGATCGAGAGCCCTCCCAACATCTCCCTGGA	1040
QY	851	CCGTCTGGAAGAACTGTGATTACTTATTAGCAGAACTCTCACTTAGAGACCGCCCAACAAAT	910
Db	1041	CCGTCTGGAAGAACTGTGATTACTTATTAGCAGAACTCTCACTTAGAGACCGCCCAACAAAT	1100
QY	911	TAAAGA/ATTGAGGAGAGCTGACGACGAAGAAGTGTCTTACAGAGGGCTTATCTATGTGACGA	970

|||||
Db 1101 TAAAGAACTGAGAGCTGACAGCAAGAAAGTCTCTACAAAGGCCAACCTATCGTCGACGA 1160
Qy 971 CCGGCCCATGCTGAGAGAGAGATCGTGACCTGTTCAGAACTTAATGAAGATGCCCTT 1030
Db 1161 CCGGCCCATGCTGAGAGAGAGATCGTGACCTGTTCAGAACTTAATGAAGATGCCCTT 1220
Qy 1031 CGTGTGAGAGGCGAGCCCTGATGCCATGACCCCGAGCGGCCCTTAGTCATCAAGAC 1090
Db 1221 CGTGTGAGAGGCGAGCCCTGATGCCATGACCCCGAGCGGCCCTTAGTCATCAAGAC 1280
Qy 1091 TGTGTCCAGTTTACCAGAAAGTCAAGTCTGTCTAAATTTTCCGAGTTGAATTATCA 1150
Db 1281 TGTGTCCAGTTTACCAGAAAGTCAAGTCTGTCTAAATTTTCCGAGTTGAATTATCA 1340
Qy 1151 GCTTAAATTTAAAGTGTGATTAAGACTCTGGGATGTGTGGCCCTCAGAGGGTC 1210
Db 1341 GCTTAAATTTAAAGTGTGATTAAGACTCTGGGATGTGTGGCCCTCAGAGGGTC 1400
Qy 1211 TGGAAATTTTAACTTGTGGCAGCAACAAAGTGAATGAATGAGAGACTTACAA 1270
Db 1401 TCGGAAATTTTAACTTGTGGCAGCAACAAAGTGAATGAATGAGAGACTTACAA 1460
Qy 1271 CCGCAGCCTGTCTGAGAGATTCAGACCTGACCTTAGGAGCAGAGATGTGGGAATCG 1330
Db 1461 CCGCAGCCTGTCTGAGAGATTCAGACCTGACCTTAGGAGCAGAGATGTGGGAATCG 1520
Qy 1331 AGGCCCTGCAATGTGATGCTGCTCTGATCGGATGAGAGCTGACCTGATCCTT 1390
Db 1521 AGGCCCTGCAATGTGATGCTGCTCTGATCGGATGAGAGCTGACCTGATCCTT 1580
Qy 1391 CGAGACTGAGAGTGTACCCAGAGGCTCTCAAGATTGACCTAGAGACCCACTCTTGGCACT 1450
Db 1581 CGAGACTGAGAGTGTACCCAGAGGCTCTCAAGATTGACCTAGAGACCCACTCTTGGCACT 1640
Qy 1451 TGTGTGATCTCCAAATCTGTCAAGTGTGCAATGCTTGGGATCAATCTGTGATTA 1510
Db 1641 TGTGTGATCTCCAAATCTGTCAAGTGTGCAATGCTTGGGATCAATCTGTGATTA 1700
Qy 1511 CATGCTGACCAATTAACCCCAAGAGAGTGAACCTTCTTCACTAAGCGCCCAATTGGAACCTG 1570
Db 1701 CATGCTGACCAATTAACCCCAAGAGAGTGAACCTTCTTCACTAAGCGCCCAATTGGAACCTG 1760
Qy 1571 GGACCAAGTGGCGAGGTGTCTGAGTGGCGGCTGTGCTCCAGCAGCAAGGAGGCTGAG 1630
Db 1761 GGACCAAGTGGCGAGGTGTCTGAGTGGCGGCTGTGCTCCAGCAGCAAGGAGGCTGAG 1820
Qy 1631 CATGAGCAGCTGACAAAGCTGTGAGAGCTCTAGGGCTGTGTGAATCTACTCAGG 1690
Db 1821 CATGAGCAGCTGACAAAGCTGTGAGAGCTCTAGGGCTGTGTGAATCTACTCAGG 1880
Qy 1691 GTGTGATCATGATGGCTAAATTTCTCAAAAGAAACATGGCTGGCAAGGGCTTCTCTT 1750
Db 1881 GTGTGATCATGATGGCTAAATTTCTCAAAAGAAACATGGCTGGCAAGGGCTTCTCTT 1940
Qy 1751 CTGGGTGTGGCTAGCAATATCATGACCTTGTGAAAAAGTATATCTTGGCCCTTTGAA 1810
Db 1941 CTGGGTGTGGCTAGCAATATCATGACCTTGTGAAAAAGTATATCTTGGCCCTTTGAA 2000
Qy 1811 TGAAGGATACATCATGTGGTTTTCATCAGCAAGAGCGGAGCGGCCCTCTAGCACAAA 1870
Db 2001 TGAAGGATACATCATGTGGTTTTCATCAGCAAGAGCGGAGCGGCCCTCTAGCACAAA 2060
Qy 1871 GCGCCCGGAGCCTTCTACTAGCGCTTCAAGGAGAGAGCAAAAGAGAGGCGTCACTTT 1930
Db 2061 GCGCCCGGAGCCTTCTACTAGCGCTTCAAGGAGAGAGCAAAAGAGAGGCGTCACTTT 2120
Qy 1931 CACTTGGTGGAAAAAGACATCAGTGGCAAGACCCAGATCCAGTCTAGAGCCATACAC 1990
Db 2121 CACTTGGTGGAAAAAGACATCAGTGGCAAGACCCAGATCCAGTCTAGAGCCATACAC 2180
Qy 1991 CAAGCAGAGCTGAACAACATGTCTATTTGGAATCATCATGGGCTATAGATCATGGA 2050
|||||

Db 2181 CAAGCAGCAGCTGAACAACATGTCTATTTGCTGAAATCATCATGGGCTATAGATCATGGA 2240
Qy 2051 TGGACCAACATCTCTGTGTCTCCACTTGTCTACCTCTACCCGACATTTCCAGAGAGA 2110
Db 2241 TGGACCAACATCTCTGTGTCTCCACTTGTCTACCTCTACCCGACATTTCCAGAGAGA 2300
Qy 2111 GGCATTTGAAAGTACTGTAGGCGCCGAGAGCCAGACACCCGCAACCCGACAGTAG 2170
Db 2301 GGCATTTGAAAGTACTGTAGGCGCCGAGAGCCAGACACCCGCAACCCGACAGTAG 2260
Qy 2171 TGGCGCCCGGACTGGAAGACCAAGTTCATCTGTGTGACACCAAGCAGCTGACCAATAC 2230
Db 2361 TGGCGCCCGGACTGGAAGACCAAGTTCATCTGTGTGACACCAAGCAGCTGACCAATAC 2420
Qy 2231 CATTGACCTGCGCATGTCTCCCGGACCTTTAGATTCATGTGATGCAAGTTTGGAAATTAACGG 2290
Db 2421 CATTGACCTGCGCATGTCTCCCGGACCTTTAGATTCATGTGATGCAAGTTTGGAAATTAACGG 2480
Qy 2291 TGAAGCTGTGAGCCCTGAGCAGAGAGAGAGAGAGCTTGTGAGTCCGCTCAAGTTTGCATGATCT 2350
Db 2481 TGAAGCTGTGAGCCCTGAGCAGAGAGAGAGAGAGAGCTTGTGAGTCCGCTCAAGTTTGCATGATCT 2540
Qy 2351 GACCTGAGAGTGTGCTACTCTCCCGCATGTGAGAGAGCTGAACCCAGAGCTGCAGAGAGCT 2410
Db 2541 GACCTGAGAGTGTGCTACTCTCCCGCATGTGAGAGAGCTGAACCCAGAGCTGCAGAGAGCT 2600
Qy 2411 GACTTGACACACCTCCCGCTGCTCCACCCCTAACCCAGCCGACCCCATATC 2462
Db 2601 GACTTGACACACCTCCCGCTGCTCCACCCCTAACCCAGCCGACCCCATATC 2652

RESULT 13
RNSTAT3PT
LOCUS RNSTAT3PT 2924 bp mRNA ROD 29-FEB-1996
DEFINITION R.norvegicus mRNA for Stat3 protein.
ACCESSION X919.0
VERSION X919.0.1 GI:1107848
KEYWORDS stat3 gene; stat3 protein.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 2924)
Ripberger, J.A., Fritz, S., Richter, K., Hocke, G.M., Lotzelsch, F. and
Fey, G.H.
Transcription factors Stat3 and Stat5b are present in rat liver
nuclei late in an acute phase response and bind interleukin-6
response elements
J. Biol. Chem. 270 (50), 29998-30006 (1995)
6102059
2 (bases 1 to 2924)
Ripberger, J.A.
Direct Submission
Submitted (26-SEP-1995) J.A. Ripberger, Universitaet
Erlangen-Nuernberg, Lehrstuhl f. Genetik, Staudtstrasse 5, D 91058
Erlangen, FRG
FEATURES
source
1..2924
Location/Qualifiers
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/dev_stage="adult"
/tissue_type="liver, acute phase"
/cell_type="predominantly hepatocytes"
/clone_lib="lambda ZapII (Stratagene)"
1..2313
/gene="Stat3"
1..2313
/gene="Stat3"
/gene="Stat3"
/note="single, polypeptide chain; needs to form homo- or
heterodimers for specific DNA-binding"
/codon_start=1
/product="Stat3 protein"

```
/protein_id="CAA62920.1"
/db_xref="GI:1107849"
/translation="MAQMOLOQIDTRYLEHOLHLYSDSEFPEMLROFLAPWIESODMA
VAASKESHATLVFNHNLGEIDDOYSRFLQSNHAAVYVTKOOLRIFLODVRKVDLEO
RIVARCLMESRLIOTATAAODGGAHNPAAVYVTKOOLRIFLODVRKVDLEO
KMKVYELQDDDFENYTKLSQGMODLNQNSVTKQKNOLEOMLTALDQMRSLY
SELAGLISAMEYOKTILDELDKRRQOJACIGEPNITCLDLRMLTISLASOLO
TRQOIKRLELOQVSYKDBIVQHRPMLERIVDLRNLMSAFVVEROPCMPHD
RPLVITKVQVFTTKVLLVPELNYOLKIVCIDKSDGVDAALRGSRRFNILQNTK
VMNEESNNGSLSEFKHLTREQRCNGGANCDAISLYTEHLITFEVEYHOGIL
KIDLETSLPYYVSNICOMPANASILMYMLNPNVNFPEKPLGTMDQYAEVL
SMOESSTKRLSTLEOLITLAEKLGSPVNSGGOTIMARKCFKPMAGKSEFVMDI
NIDLYKXILALMBGYIMGFISKEPRALTSKPPGTLLRSESSKBGVPTFY
EKDISKTOIOSVEPTKQOLNNMSFAEIIINGVITMDNTNLVSPLYLIPDIKEA
FGKYCRPESQHPADPQSAAPYLKTKFICVTPPTYSNTIDLPMSPTLSDLMQFNN
GEQAPAGQOFESLTFDMDLTSECATSPM"
```

BASE COUNT 771 a 732 c 784 g 637 t
ORIGIN

```
Query Match 80.0% Score 2296; DB 10; Length 2924;  
Best Local Similarity 90.8%; Pred. No. 0;  
Matches 2550; Conservative 0; Mismatches 210; Indels 48; Gaps 8;
```

OY 69 ATGGCTCAGTGGACACCTGCAGCGCTGGACACACGCTTACCTGAGCAGCTGCAGCAG 128
DB 1 ATGGCTCAGTGGACACCTGCAGCGCTGGACACACGCTTACCTGAGCAGCTTATCAG 60

OY 129 CTGTACAGCAGACACCTGCCATGTGCGGAGCTTCTGTGACCTGTGATGAGAGT 188
DB 61 CTGTACAGCAGTGTGCTCCCATGTGCGGAGCTTCTGTGACCTGTGATGAGAGC 120

OY 189 CAGACCTGGGCTATGACCGCAGCAAGATGACATGCCAGCTTGTGTTATATCTC 248
DB 121 CAGACCTGGGCTATGACCGCAGCAAGATGACATGCCAGCTTGTGTTATATCTC 180

OY 249 TTGGGTGAATTTGACCCACCAATATAGCCGATTCCTGCAAGATGCCAATGCTCTATCAG 308
DB 181 TTGGGTGAATTTGACCCACCAATATAGCCGATTCCTGCAAGATGCCAATGCTCTATCAG 240

OY 309 CACAACCTTTCAGAAATCAGCAGTTTTCGACAGCAGGATCTTGAGAACCAATGGAA 368
DB 241 CACAACCTTTCAGAAATCAGCAGTTTTCGACAGCAGGATCTTGAGAACCAATGGAA 300

OY 369 ATTGCCGATGCTGCCCGATGCTGTGGGAAGAGTCTGCCGCTCTCCAGACGGCAGCC 428
DB 301 ATTGCCGATGCTGCCCGATGCTGTGGGAAGAGTCTGCCGCTCTCCAGACGGCAGCC 360

OY 429 ACGCAGCCCGACAGAGGGGCCAGGCCAACCAACCAAGCCGCGTAGTGACAGAGAAG 488
DB 361 ACGCAGCCCGACAGAGGGGCCAGGCCAACCAACCAAGCTGCGTAGTGACAGAGAAG 420

OY 489 CAGCAGATGTTGGACACATCTTACAGATGTCGGAGCGAGTGCAGAGATTAACAACAG 548
DB 421 CAGCAGATGTTGGACACATCTTACAGATGTCGGAGCGAGTGCAGAGATTAACAACAG 480

OY 549 AAAATGAGGTGGTGGAGAACCTCCAGACGACTTTGATTTCACATACAAAACCTTCAAG 608
DB 481 AAAATGAGGTGGTGGAGAACCTCCAGAGTGTTCGATTTCACATACAAAACCTTCAAG 540

OY 609 AGCCAGAGGACATCGAGATTTGAAATGGAAACCAACCACTCTGTACACAGACAGAAGATG 668
DB 541 AGTCAAGAGACATCGAGATTTGAAATGGAAACCAACCACTCTGTACACAGACAGAAGATG 600

OY 669 CAGCAGCTGGAACAGATCTGCACAGCTGGACAGATGGGAGAGAACTTGTGATGAG 728
DB 601 CAGCAGCTGGAACAGATCTGCACAGCTGGACAGATGGGAGAGAACTTGTGATGAG 660

OY 729 CTGGCGGGGCTCTTGTGACGCAATGAGTACGAGAAAGACACTGATGATGAGAGCTG 788
DB 661 CTGGCGGGGCTCTTGTGACGCAATGAGTACGAGAAAGACACTGATGATGAGAGCTG 720

OY 789 GCTGACTGGAAGAGGGCGCCAGACATCCCGTGCATCGGAGGGCCCTCCCAACATCTGCTG 848
DB 721 GCTGACTGGAAGAGGGCGCCAGACATCCCGTGCATCGGAGGGCCCTCCCAACATCTGCTG 780

OY 849 GACCCCTGGAAGAACTGATTAACCTTATACAGAAATCTCAACTTCAAGCCGCCAACAA 908
DB 781 GACCCCTGGAAGAACTGATTAACCTTATACAGAAATCTCAACTTCAAGCCGCCAACAA 840

OY 909 ATTAAGAACTGAGAGAGCTGACGAGAAAGTCTCTCAAGGGCGACCTTATGTCAG 966
DB 841 ATTAAGAACTGAGAGAGCTGACGAGAAAGTCTCTCAAGGGCGACCTTATGTCAG 900

OY 969 CACGGGCCATGCTGGAGAGAGAGTGTGGAGCTGTGCAAGAACTTAATAAGAGTCC 1027
DB 901 CACGGGCCATGCTGGAGAGAGAGTGTGGAGCTGTGCAAGAACTTAATAAGAGTCC 960

OY 1029 TTCTGCTGAGAGCGGCGACCCCTGATGCCATGCAACCCGAGCCGCTTATGATCAAG 1088
DB 961 TTCTGCTGAGAGCGGCGACCCCTGATGCCATGCAACCCGAGCCGCTTATGATCAAG 1020

OY 1089 ACTGCTGTCAGCTTATACCAAGAACTCAGGTTGCTGCTCAAAATTTCTGAGTGAATAT 1148
DB 1021 ACTGCTGTCAGCTTATACCAAGAACTCAGGTTGCTGCTCAAAATTTCTGAGTGAATAT 1080

OY 1149 CAGCTTAAATTAAGTGTGATTAAGACCTGTGGGATGTGCTGCTGAGAGG 1208
DB 1081 CAGCTTAAATTAAGTGTGATTAAGACCTGTGGGATGTGCTGCTGAGAGG 1140

OY 1209 TCTGGAAATTTAACATTTCTGGGCGAAGACACAAAAGTGATGAACATGAGAGCTTAAC 1268
DB 1141 TCTGGAAATTTAACATTTCTGGGCGAAGACACAAAAGTGATGAACATGAGAGCTTAAC 1200

OY 1269 AAGCGACCTGCTGCGACAGTTCAGACCTGACCTGACCTTACGAGACAGAGTGGGAT 1328
DB 1201 AAGCGACCTGCTGCGACAGTTCAGACCTGACCTGACCTTACGAGACAGAGTGGGAT 1260

OY 1329 GGAGGCCGTGCCAATTTGATGATCCCTCTTGTATGCTGACTGAGAGAGCTCAGCTGATCAC 1388
DB 1261 GGAGGCCGTGCCAATTTGATGATCCCTCTTGTATGCTGACTGAGAGAGCTCAGCTGATCAC 1320

OY 1389 TTGAGACTGAGGTGTACCCAGAGGCTCAAGATTTGACTAGAGACCCACTCTTGCCA 1448
DB 1321 TTGAGACTGAGGTGTACCCAGAGGCTCAAGATTTGACTAGAGACCCACTCTTGCCA 1380

OY 1449 GTTGTGATGATCTCCAAATCTGTGAGATGCCAAATGCTGGGCTATCAATCTGTGAT 1508
DB 1381 GTTGTGATGATCTCCAAATCTGTGAGATGCCAAATGCTGGGCTATCAATCTGTGAT 1440

OY 1509 AACATGCTGACCAATTAACCCCAAGAACTGAATCTTCTACTAAGCGCCCAATTGGAAC 1568
DB 1441 AACATGCTGACCAATTAACCCCAAGAACTGAATCTTCTACTAAGCGCCCAATTGGAAC 1500

OY 1569 TTGGACCAAGTGGCGGAGTGTCTGAGCTGGCAAGTCTGTCCACCAACAGAGGGCTG 1628
DB 1501 TTGGACCAAGTGGCGGAGTGTCTGAGCTGGCAAGTCTGTCCACCAACAGAGGGCTG 1560

OY 1629 AGCATCGAGCAGCTGACAAACGCTGCTGAGAAAGCTCTAGGGCTGTGTAACTACTCA 1688
DB 1561 AGCATCGAGCAGCTGACAAACGCTGCTGAGAAAGCTCTAGGGCTGTGTAACTACTCA 1620

OY 1689 GGGTGTCAAGTACATCTGGGCTTAATTTGCAAAAGAAACATGCTGGCAGAGGGCTTCTCC 1748
DB 1621 GGGTGTCAAGTACATCTGGGCTTAATTTGCAAAAGAAACATGCTGGCAGAGGGCTTCTCC 1680

OY 1749 TTCTGGTCTGGCTGAGCAATATCATGACCTGTGAAAGAGTATATTTGGCCCTTTGG 1808
DB 1681 TTCTGGTCTGGCTGAGCAATATCATGACCTGTGAAAGAGTATATTTGGCCCTTTGG 1740

OY 1809 AATGAAGGCTATCATCATGAGTTTATCATGCAAGAGCGGGAGCGGCGCATCTAAGACA 1868
DB 1741 AATGAAGGCTATCATCATGAGTTTATCATGCAAGAGCGGGAGCGGCGCATCTAAGACA 1800

OY 1869 AAGCCCCCGGCGACCTTCTCTACTGCGCTTCAAGCGAGACACCAAGAAAGAGGGGTCACT 1928

Db	1801	AAGCCCCGGGACCTTCCGTGCGGGTTCACGTGAGAGACGACGAAGGAGGGGTACT	1860
Oy	1929	TTCACTTGGGTGAAAGAGACATCACTGGCAAGACCAGATTCACCTCTGTAGAGCCATAC	1988
Db	1861	TTCACTTGGGTGAAAGAGACATCACTGGCAAGACCAGATTCACCTCTGTAGAACCAAT	1920
Oy	1989	ACCAAGCAGAGAGTGAACAACATGATTTGGTGAATATCATGGGTATTAAGATCATG	2048
Db	1921	ACCAAGCAGAGAGTGAACAACATGATTTGGTGAATATCATGGGTATTAAGATCATG	1980
Oy	2049	GATGCGACCAACATCTGTGTCTCCACTTGTCTACCTTACCCGACATTTCCAAAGAG	2108
Db	1981	GACCGTACCAACATCTGTGTATCCCACTGGTCTACCTTACCCGACATTTCCAAAGAG	2040
Oy	2109	GAGGCATTTGGAAAGTACTGTACGGCCGAGAGCAGACACCCTGGAAGCCGACCAGT	2168
Db	2041	GAGGCATTTGGAAAGTATGTTCGCCCCGAGAGCCAGAGACACCTTGAAGCTGACCAGT	2100
Oy	2169	AGTGTCTCCCCGTACTGGAAGACCAAGTTTCATCTGTGTACACCAAGACCTGTACGAT	2228
Db	2101	AGTGTCTCCCCCTTACCTGAAGACCAAGTTTCATCTGTGTACACCAAGACCTGTACGAT	2160
Oy	2229	ACCATTTACCTGCCCATGTGTCCCCCGCCACTTTAGATTCATATATGACATTTGGAAATAC	2288
Db	2161	ACCATTTACCTGCCCATGTGTCCCCCGCCACTTTAGATTCATATATGACATTTGGAAATAC	2220
Oy	2289	GGTAAAGGTCTGAGCCCTCAGCAGAGAGGAGGACATTTGATGCTGCTACGTTTGACATGAT	2348
Db	2221	GGTAAAGGTCTGAGCCCTCAGCAGAGAGGAGGACATTTGATGCTGCTACGTTTGACATGAT	2280
Oy	2349	CTGACCTCGGAGTGTGCTACCTCCCCCATGTGAGAGCTGAAACCAAGACTGCAGAGAC	2408
Db	2281	CTGACCTCGGAGTGTGCTACCTCCCCCATGTGAGAGCTGAAAGGCAAGACTGCAGAG--	2338
Oy	2409	GTGACTTGAGACACCTCCCCGTTGCTCACCCCTAAGACGCGGAACCCCATATCTCTGA	2468
Db	2339	-----TATGACTGAGACACTGCCCATGTACCTTAGCAG-----CAACTCA	2379
Oy	2469	AACCTCTAACCTTGTGGTTCCAGATTTTTTTTTTAAATTTCTACTCTGTACTTCTTGG	2528
Db	2380	GATCACTGAACACTACATGTGTCCAGATTTTTTTTTTAACTCTCACTGCTGCTTGG	2439
Oy	2529	GCATCTGGGCACTTTTAAAGAGAAATAGTAGTGGGTGGATAACTGTTAGT	2588
Db	2440	GCATCTGGGCACTTTTAAATAGAGAACTGATGAGTGGGTGGATGCTGTTAGT	2499
Oy	2589	AAAGAGAGAG--ACCTCTGAGTCTGGGGATGGGGCTGACAGAGAGA-----AGGAGGCA	2641
Db	2500	GAGGAGGGGGCACTCTTGAATGCCGATGCGTGTGAAGAGAGAGGGCTGGGGGGAGC	2559
Oy	2642	AAGGGGAACACCTCTGTCTCTGCCCGGCTGCTCCTTTTTCAGCAGCTGGGGGGTGT	2701
Db	2560	GAAAGGAACACCTATATTTCTGTCCGCTGTCTCTCT-----GGTTTG	2602
Oy	2702	TGTTAGACAAGTGCCTCTGTGGCCATGAGCTACCTGTGGCCCACTGTGAGCTGATA	2761
Db	2603	TTGTTGACAACGCTCTCTGTGT--CCATGGCATCTGCTGCTCCCTTTCTGTAGCTGATA	2661
Oy	2762	CCCCATTTGGGAACTCTGGCTCTGCACATTTCAACCTTGCTAATATCCATATAGAGCT	2821
Db	2662	CCCCAGGCTGG--ACTCCGGCTCTGCACATTTCAACCTGTTAATGTCCATATAGAGCT	2720
Oy	2822	AGGACTAAAGCCAGAGAGTCTCTTAAATTAAATTAAGAAAAAAAAAAAAA	2869
Db	2721	AGGACTAAAGCCTA--TAGGCTTCTTTAACTTAAATTAAGAAAAAAAAAAAAATA	2767

RESULT	14			
188782				
LOCUS	188782	2310 bp	DNA	PAT
DEFINITION	Sequence 6 from patent US 5719042.			10-AUG-1998
ACCESSION	188782			

VERSION	188782.1	GI:3408722
KEYWORDS	..	
SOURCE	unknown	
ORGANISM	unknown	
REFERENCE	Unclassified.	
AUTHORS	1 (bases 1 to 2310)	
TITLE	Kishimoto, T. and Akira, S.	
JOURNAL	Nucleic acids encoding transcription factor APRE (acute phase response factor)	
FEATURES	Patent: US 5719042-A 6 17-FEB-1998;	
SOURCE	Location/Qualifiers	
	1..2310	
BASE COUNT	615 a 597 c 629 g 469 t	
ORIGIN	1..2310	
Query Match	80.0%; Score 2294; DB 6; Length 2310;	
Best Local Similarity	99.6%; Pred. No. 0;	
Matches 2300; conservative	0; Mismatches 10; Indels 0; Gaps 0;	
QY	ATGGCTCACTGGACACCACTGTCAGACAGCTGGACACACGCTACTTGAAGCAGCTGCACCG	128
DB	1 ATGGCTCACTGGACACCACTGTCAGACAGCTGCAGACACACGCTACTTGAAGCAGCTGCACCG	60
QY	129 CTGTACAGCGACAGCTTCCCATGTGAGCTGGCGGAGTTCTCGACCTTGGATTGAGAGT	188
DB	61 CTGTACAGCGACAGCTTCCCATGTGAGCTGGCGGAGTTCTCGACCTTGGATTGAGAGT	120
QY	189 CAAGACTGGGCAATATGACAGCCAGCAAAAGAGTACATGCCAGCTGTGTTATATATCTC	248
DB	121 CAAGACTGGGCAATATGACAGCCAGCAAAAGAGTACATGCCAGCTGTGTTATATATCTC	180
QY	249 TTGGGTGAATTTGACCCAGCAATATAGCCGATTTCTGCAAGAGTCCAAATGTCCTCTATGAG	308
DB	181 TTGGGTGAATTTGACCCAGCAATATAGCCGATTTCTGCAAGAGTCCAAATGTCCTCTATGAG	240
QY	309 CACAACTTCGAAGAATCAAGCAGTTTCTGCGAGAGCAGGTATCTTGAGAGCCATGGAA	368
DB	241 CACAACTTCGAAGAATCAAGCAGTTTCTGCGAGAGCAGGTATCTTGAGAGCCATGGAA	300
QY	369 ATTGCCCGGATGCTGGCCCGATGCCCTGTGGGAAGAGTCTCGCTCCAGACGGCAGCC	428
DB	301 ATTGCCCGGATGCTGGCCCGATGCCCTGTGGGAAGAGTCTCGCTCCAGACGGCAGCC	360
QY	429 ACGGCACGCCACGAAGGGGCCAGGCCAACCCACCAACAGCCGCGGTAGTACAGAGAAG	488
DB	361 ACGGCACGCCACGAAGGGGCCAGGCCAACCCACCAACAGCCGCGGTAGTACAGAGAAG	420
QY	489 CAGCAGATTTGGAGCAGCATCTTCAGATGTCGGAAACGAGTCCAGAGTATAGAACAG	548
DB	421 CAGCAGATTTGGAGCAGCATCTTCAGATGTCGGAAACGAGTCCAGAGTATAGAACAG	480
QY	549 AAAATGAGGTGTGGAGAACCTCCAGACGACCTTTGATTTCACTACAAAACCTCAAG	608
DB	481 AAAATGAGGTGTGGAGAACCTCCAGACGACCTTTGATTTCACTACAAAACCTCAAG	540
QY	609 ACGCAAGGAGCATGTCAGATCTGAAATGGAACAACCACTGTGTGACCCAGACAGAGAAGT	668
DB	541 ACGCAAGGAGCATGTCAGATCTGAAATGGAACAACCACTGTGTGACCCAGACAGAGAAGT	600
QY	669 CAGCAGCTGGAACAGATGCTCACAGCCCTGGAACAGATGCGAGAGAAGCATTTGTGATGAG	728
DB	601 CAGCAGCTGGAACAGATGCTCACAGCCCTGGAACAGATGCGAGAGAAGCATTTGTGATGAG	660
QY	729 CTGGCGGGGCTCTTGTACACCATGGAGTACGAGCAAGAACACTGATGATGAAGAGCTG	788
DB	661 CTGGCGGGGCTCTTGTACACCATGGAGTACGAGCAAGAACACTGATGATGAAGAGCTG	720
QY	789 GCTGACTGGAAGAGCGGCGCAGAGATCGGCTGATCGAGAGGCGCTCCCAACATCTGCCTG	848
DB	721 GCTGACTGGAAGAGCGGCGCAGAGATCGGCTGATCGAGAGGCGCTCCCAACATCTGCCTG	780

QY	849	GACCCTCTGAAAACTGGATAACTTCATTAGCAAGATCTCAACTTCAGACCCCGCCAAACA	908
Db	781	GACCCTCTGAAAACTGGATAACTTCATTAGCAAGATCTCAACTTCAGACCCCGCCAAACA	840
QY	909	ATTAGAAACTGGAGGAGCTGCAGCAAGAAAGTGCCTCAAAAGGAGACCTATGTCGAC	968
Db	841	ATTAGAAACTGGAGGAGCTGCAGCAAGAAAGTGCCTCAAAAGGAGACCTATGTCGAC	900
QY	969	CACCGGCCCATGCTGGAGAGAGAGATCGTGGAGCTGTCAGAACTTAATGAAGTGGC	1028
Db	901	CACCGGCCCATGCTGGAGAGAGATCGTGGAGCTGTCAGAACTTAATGAAGTGGC	960
QY	1029	TTCTGGTGGAGCGCGCACCCCTCGATGCCATGCAACCCGAGACGGGCGCTTGTGATCAAG	1088
Db	961	TTCTGGTGGAGCGCGCACCCCTCGATGCCATGCAACCCGAGACGGGCGCTTGTGATCAAG	1020
QY	1089	ACTGCTGTCCAGTTTACACAGAAAGTCAGTGTGCTGGTCAAAATTTCCAGATTGAATTAT	1148
Db	1021	ACTGCTGTCCAGTTTACACAGAAAGTCAGTGTGCTGGTCAAAATTTCCAGATTGAATTAT	1080
QY	1149	CAGCTTAAATTTAAAGTGTGCTATTGATAAAAGACTCTGGGGATGTTGTCTGCTCAGAGG	1208
Db	1081	CAGCTTAAATTTAAAGTGTGCTATTGATAAAAGACTCTGGGGATGTTGTCTCAGAGG	1140
QY	1209	TCTCGGAAATTTAACTTCTGGGCACAGAACACAAAAGTATACATGAGAGAGCTTAAAC	1268
Db	1141	TCTCGGAAATTTAACTTCTGGGCACAGAACACAAAAGTATATTAACATGAGAGAGCTTAAAC	1200
QY	1269	AACGGCAGCCTGTCTGCGAGAGTTCCAAACACTGACCTTAGCGAGACAGAGATGTGGCAAT	1328
Db	1201	AACGGCAGCCTGTCTGCGAGAGTTCCAAACACTGACCTTAGCGAGACAGAGATGTGGCAAT	1260
QY	1329	GGAGCGCGTGGCAATTTGTGATGGCTCCCTCTATGTACTGAGAGAGTGGACTGTATCAAC	1388
Db	1261	GGAGCGCGTGGCAATTTGTGATGGCTCCCTCTATGTACTGAGAGAGTGGACTGTATCAAC	1320
QY	1389	TTTGGAGTGAAGTGTATGCACCAAGGCGCTCAAGATTGACCTTAGAGACCCCACTCTTTGCCA	1448
Db	1321	TTTGGAGTGAAGTGTATGCACCAAGGCGCTCAAGATTGACCTTAGAGACCCCACTCTTTGCCA	1380
QY	1449	GTTGGGTGATCTCCAACTCTGTGCAAGATGCCAAATGCTTTGGGCATCAATCCTGTGGTAT	1508
Db	1381	GTTGGGTGATCTCCAACTCTGTGCAAGATGCCAAATGCTTTGGGCATCAATCCTGTGGTAT	1440
QY	1509	AACATGCTGACCAATTAACCCCAAGAAAGTGAACCTTTCATCTAAAGCCGCCAATTTGGAAC	1568
Db	1441	AACATGCTGACCAATTAACCCCAAGAAAGTGAACCTTTCATCTAAAGCCGCCAATTTGGAAC	1500
QY	1569	TGGGACCAAGTGGCGGAGGTGCTCAGTGGCAATTTCTGTCCACCAACCAAGCGAGGCGTG	1628
Db	1501	TGGGACCAAGTGGCGGAGGTGCTCAGTGGCAATTTCTGTCCACCAACCAAGCGAGGCGTG	1560
QY	1629	AGCATCGAGCAGCTGAGACAAGCGTGGCTGGAACAACTCTTAGGGGCTGTGTTGAACACTCA	1688
Db	1561	AGCATCGAGCAGCTGAGACAAGCGTGGCTGGAACAACTCTTAGGGGCTGTGTTGAACACTCA	1620
QY	1689	GGGTGTCAAGATCAATGSGGCTAAATTTGTCAAGAAACAATGGCTGGCAAGGGCTTCTACC	1748
Db	1621	GGGTGTCAAGATCAATGSGGCTAAATTTGTCAAGAAACAATGGCTGGCAAGGGCTTCTACC	1680
QY	1749	TTTGGGCGTCTGGCTAGACAAATATCATGACCTTGTAAGAAAGATATCTTGGCCCTTTGG	1808
Db	1681	TTTGGGCGTCTGGCTAGACAAATATCATGACCTTGTAAGAAAGATATCTTGGCCCTTTGG	1740
QY	1809	AATGAAGGATACATCATGCGTTTCATAGCAAGAGACGGGAGCGGCCATCTTAAGACA	1868
Db	1741	AATGAAGGATACATCATGCGTTTCATAGCAAGAGACGGGAGCGGCCATCTTAAGACA	1800
QY	1869	AAGCGCCCGGGCACCCTTCCTACTGCGCTTAGGAGAGCAGCAAGAAAGAAAGAGGGTCACT	1928
Db	1801	AAGCGCCCGGGCACCCTTCCTACTGCGCTTAGGAGAGCAGCAAGAAAGAAAGAGGGTCACT	1860
QY	1929	TTTCACTTGGGTGAAAGAGCATCACTGTGCAAGACCCAGATCCAGTCTGTAGACCATATAC	1988

Db	1861	TTCACTTGGGTGGAAAGGACATCACTGGCGAGACCAGATCCAGTCTGTAGGCAATAC	1920b
Qy	1989	ACCAAGCAGCAGCTGGAACAACATGTCATTTGCTGAATCATGGCTATAGATCATG	2048b
Db	1921	ACCAAGCAGCAGCTGAGCAACATGTCACTTTGCTGAATCATGGCTATAGATCATG	1980b
Qy	2049	GATGCGACCAACATCTGTGTGTCTCCACTTGTCTACCTTACCCGCACTTCCCAAGAG	2108b
Db	1981	GATGCGACCAACATCTGTGTGTCTCCACTTGTCTACCTTACCCGCACTTCCCAAGAG	2040b
Qy	2109	GAGGATTTGGAAGTACTGTAGCGCCGAGACGAGAGACACCCGAGACCAACCAAGT	2168b
Db	2041	GAGGATTTGGAAGTACTGTAGCGCCGAGAGACGAGAGACACCCGAGAGCAACCAAGT	2100b
Qy	2169	AGTGTGCCCCGTACCTGAAGACCAAGTTATCTGTGTGACACCAAGCACTGCAGCAAT	2228b
Db	2101	AGTGTGTCTGCTACCTGAAGACCAAGTTATCTGTGTGACACCAAGCACTGCAGCAAT	2160b
Qy	2229	ACCATTGACCTGCCGATGTCGCCCGGCACTTATGATTCATTGATGCACTTTGGAATTAAC	2288b
Db	2161	ACCATTGACCTGCCGATGTCGCCCGGCACTTATGATTCATTGATGCACTTTGGAATTAAC	2220b
Qy	2289	GGTGAAGTGTCTGAGCCCTCAGCAGAGGGCACTTTGAGTGCCTACGTTTACATGGAT	2348b
Db	2221	GGTGAAGTGTCTGAGCCCTCAGCAGAGGGCACTTTGAGTGCCTACGTTTACATGGAT	2280b
Qy	2349	CTGACCTCGAGTGTGCTACCTCCCATG	2378b
Db	2281	CTGACCTCGAGTGTGCTACCTCCCATG	2310b

RESULT	15						16-MAY-2001
LOCUS	AR120980						
DEFINITION	Sequence 1 from patent US 6159694.						
ACCESSION	AR120980						
VERSION	AR120980.1						
KEYWORDS	GI:14104556						
SOURCE	Unknown.						
ORGANISM	Unclassified.						
REFERENCE	1 (bases 1 to 2787)						
AUTHORS	Karras,J.G.						
TITLE	Antisense modulation of stat3 expression						
JOURNAL	Patent: US 6159694-A 1 12-DEC-2000;						
FEATURES	Location/Qualifiers						
source	1..2787						
	/organism="unknown"						
BASE COUNT	729 a	719 c	586 t				
ORIGIN							

[illegible]

```
OY 301 TCTATCAGCACACCTTTCGAGAAATCAAGCACTTCTCGACAGGAGGATCTGTGAGAC 360
    |||||
Db 453 TCTATCAGCACATCTAGGAAATCAAGCACTTCTTCAGAGCAGGTATCTTGAGAAC 512
OY 361 CAATGGAAATTCGCCGATCGTGGCCCGATGCTGTGGGAAAGTCTGCTCTCCAGA 420
    |||||
Db 513 CAATGGAAATTCGCCGATGTGTGCCCGTGTGCTGTGGGAAAGTCAAGCCCTTTACAGA 572
OY 421 CGGACGCCAGCCAGCCAGCAAGGGGGCCAGGCCAACCAACCAAGCCGCCAGTGA 480
    |||||
Db 573 CTGACGCCACTCGCGGCCAGCAAGGGGGCCAGGCCAACCAACCAAGCCAGTGTGA 632
OY 481 CAGAGAGCAGCATGTTGAGAGCAGCATCTCAGATGTCCGGAAGCGAGCGAGATC 540
    |||||
Db 633 CGAGAGAGCAGCATGATGCTGTGAGAGCAGCATCTCAGATGTCCGGAAGAGAGCGAGATC 692
OY 541 TAGAACGAAATGAAAGTGTGTGAGACCTCCAGAGCAGCTTTGATTTCAATACAAA 600
    |||||
Db 693 TAGAACGAAATGAAAGTGTGTGAGAAATCTCCAGGATGACTTTGATTTCAACTATMAA 752
OY 601 CCTCAGAGCCAGGAGAGCATGAGATCTGAAATGAAACCAACGATCTGTGACAGAC 660
    |||||
Db 753 CCTCAGAGGTCAAGAGAGCATGCAAGATCTGAAATGAAACCAACGATGAGACAGGC 812
OY 661 AGAAGATGAGCAGCATGGAACAGATGCTCAGACCCCTGAGCAGATGCGGAGAGCATTTG 720
    |||||
Db 813 AGAAGATGAGCAGCATGGAACAGATGCTCAGCTGAGCAGCATGAGAGAGCATG 872
OY 721 TCGATGAGCTGGGGGGCTTGTTCAGCAATGAGATGAGTGCAGAGAGACACTGACGATG 780
    |||||
Db 873 TCGATGAGCTGGGGGGCTTGTTCAGCGATGAGATGAGTGCAGAGAACTCTCAGGAGC 932
OY 781 AAGAGCTGCTGACTGGAAGAGCGCCAGAGATCGCTGATCGAGAGCCCTCCACAGA 840
    |||||
Db 933 AGAGAGCTGGTGACTGGAAGAGCGCCAGATGAGTGCATTTGAGAGCGCCCCACAGA 992
OY 841 TCTGCTGAGCCGCTGTGAGAAATCGATTAATCTATTAGAGAGATGCAACTTCAACCC 900
    |||||
Db 993 TCTGCTGAGTGTGAGAAATCGATTAATCTATTAGAGAGATGCAACTTCAACCC 1052
OY 901 GCCAACAAATTAAGAAATCGAGAGCTGAGAGAAAGTGTCTCAAGGGCGAGCCTTA 960
    |||||
Db 1053 GTCAACAATTAAGAAATCGAGAGAGTTCACCAAAAAGTTCTCTCAAAAAGGGAGCCCA 1112
OY 961 TCGTGCAGCACCGCCCATCTGTGAGAGAGAGATGCTGAGCTGTTCAGAACTTAATGA 1020
    |||||
Db 1113 TTGTACAGCACCGCCCATCTGTGAGAGAGAGATCGTGTGAGCTGTTCAGAACTTAATGA 1172
OY 1021 AGAGTGCCTTCGTGTGAGAGCGGAGCCCGCATGCCATGCCAGCACCGCGGAGCCCTTAG 1080
    |||||
Db 1173 AAGGTGCTTTGTGTGTGAGAGCGGAGCCCGCATGCCATGCCAGTGCAGCCGGCCCTCG 1232
OY 1081 TCATCAGAGCTGTGTGCTCAGTTTACAGAAAGTCAAGTGTGTCTCAAAATTTCTGAGT 1140
    |||||
Db 1233 TCATCAGAGACGGGGTCTCAGTTTCACTAAAGTCAAGTGTGTCTCAAGTTCCCTGAGT 1292
OY 1141 TGAATTTATCAGCTTTAAATTTAAAGTGTGATTAAGAGCTCTGGGAGATGTGTGCTCC 1200
    |||||
Db 1293 TGAATTTATCAGCTTTAAATTTAAAGTGTGATTAAGAGAGCTCTGGGAGAGTGTGACCTC 1352
OY 1201 TCAGAGGCTCTCGGAAATTTAACTTTCTGGGACAGACACAAAAAGTATGAATGAGAG 1260
    |||||
Db 1353 TCAGAGGATCCCGGAAATTTTACATTTCTGGGACAAACACAAAAAGTATGAACATGGAAG 1412
OY 1261 AGTCTAACACGCGAGCTGTCTGTGAGAGTTCAAGCAGCTGACCTTATAGGAGCAGAGAT 1320
    |||||
Db 1413 AATCTAACACGCGAGCTGTCTGTGAGAGTTCAACACCTTGAAGAGAGAGAGAT 1472
OY 1321 GTGGGAATGAGAGCGGTGTGATGATGCTCTTGTGATGAGTGTGAGAGAGCTGACAC 1380
    |||||
Db 1473 GTGGGAATGAGAGCGGTGTGATGATGCTCTTGTGATGAGTGTGAGAGAGCTGACAC 1532
OY 1381 TGATCACCCTTGAGAGCTGAGGTGTACCAACCAAGGCTCAAGATTGACCTAGAGACCACT 1440
    |||||

Db 1533 TGATCACCCTTGAGAGCCGAGGTGTATCACCAAGGCTCAAGATTGACCTAGAGACCACT 1592
OY 1441 CCTTGCCAGTTGTGTGATCTTCACCAATCTGTGAGATGCAAAATGCTTGGGCAATCTC 1500
    |||||
Db 1593 CCTTGTCAAGTTGTGTGATCTTCACCAATCTGTGAGATGCAAAATGCTTGGGCAATCTC 1562
OY 1501 TGTGTATTAACATGCTGTGACCAATAACCCCAAGAACGTGAACCTTCTTCACTAAGCCCAA 1560
    |||||
Db 1653 TGTGTATTAACATGCTGTGACCAATAACCCCAAGAACGTGAACCTTCTTCACTAAGCCCAA 1712
OY 1561 TTGGAACCTGGGACCAAGTGGCCGAGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1620
    |||||
Db 1713 TTGGAACCTGGGACCAAGTGGCCGAGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1772
OY 1621 GAGGCTGAGCATGCGAGCAGTGTGACCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1680
    |||||
Db 1773 GAGGCTGAGCATGCGAGCAGTGTGACCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1832
OY 1681 ACTACTCAGGCTGTGAGATGACATGAGCTTAATTTCTGCAAAACAAACATGCTGTGAGAG 1740
    |||||
Db 1833 ACTACTCAGGCTGTGAGATGACATGAGCTTAATTTCTGCAAAACAAACATGCTGTGAGAG 1892
OY 1741 GCTTGTCTTGTGGGTGTGGGTGTGAGCAATATGATGACCTTGTGAAAAAGTATATCTGG 1800
    |||||
Db 1893 GCTTGTCTTGTGGGTGTGGGTGTGAGCAATATGATGACCTTGTGAAAAAGTATATCTGG 1952
OY 1801 CCTTTTGAATGAAGGTATACATCATGAGTGTTCATCAGCAAGAGCGGGAGCGGCATTC 1860
    |||||
Db 1953 CCTTTTGAATGAAGGTATACATCATGAGTGTTCATCAGCAAGAGCGGGAGCGGCATTC 2012
OY 1861 TAAACCAAAAGCCCCGGGACCTTCTCTAGCGCTTCAGCGAGAGCAACAAAGAGAG 1920
    |||||
Db 2013 TAGCACTAAAGCCCCGGGACCTTCTCTAGCGCTTCAGCGAGAGCAACAAAGAGAG 2072
OY 1921 GGGTCACTTTCATGTTGGGTGTGAAAAAGGACATGAGTGTGCAAGAGCCAGATCCAGTCTGAG 1980
    |||||
Db 2073 GGGTCACTTTCATGTTGGGTGTGAAAAAGGACATGAGTGTGCAAGAGCCAGATCCAGTCTGAG 2132
OY 1981 AGCCATACCAAGCAGCAGCTGGAACACATGTCATTTCTGAAATCATATGAGCTATA 2040
    |||||
Db 2133 AACCATACCAAGCAGCAGCTGGAACACATGTCATTTCTGAAATCATATGAGCTATA 2192
OY 2041 AGATCATGATGATGAGCAACATCTGTGTCTGACCTTGTGATGATGATGATGATGATG 2100
    |||||
Db 2193 AGATCATGATGATGAGCAACATCTGTGTCTGACCTTGTGATGATGATGATGATGATG 2252
OY 2101 CCAAGAGAGAGCATTTGAAAGTACTGTAGGCGCCGAGAGCCAGAGACCCCGAAGCGG 2160
    |||||
Db 2253 CCAAGAGAGAGCATTTGAAAGTACTGTAGGCGCCGAGAGCCAGAGATTCCTGAAAGCTG 2312
OY 2161 ACCCAGTATGCTGCCCCGTACCTCAAGACCAAGTTTCATGTGTGACACCAAGACCT 2220
    |||||
Db 2313 ACCCAGTATGCTGCCCCGTACCTCAAGACCAAGTTTCATGTGTGACACCAAGACCT 2372
OY 2221 GCAGCAATACATTTGACCTGCGATGTCGCCCGCACTTGTAGATGATGATGATGATGATG 2280
    |||||
Db 2373 GCAGCAATACATTTGACCTGCGATGTCGCCCGCACTTGTAGATGATGATGATGATGATG 2432
OY 2281 GAAATTAAGGTGAAGGTGCTGAGCCCTCAGACGAGAGGCAAGTTTGTAGTGTGCTACGTTTG 2340
    |||||
Db 2433 GAAATTAAGGTGAAGGTGCTGAGCCCTCAGACGAGAGGCAAGTTTGTAGTGTGCTACGTTTG 2492
OY 2341 ACATGATGATGACCTGCGAGTGTGTACCTCCCGCATGTGAGAGAGTGTGAACACGAGAGT 2400
    |||||
Db 2493 ACATGATGATGACCTGCGAGTGTGTACCTCCCGCATGTGAGAGAGTGTGAACACGAGAGT 2552
OY 2401 GC----AGAGCACTGACTTGTGAGACACTGCGCGCTCTCCACCCCTTAACACCAAGCC 2456
    |||||
Db 2553 GCAGAAAGATAGACAGTGAAGCGGCTCTACCTGATTTGCGCACCCCTCAGACAGCAAAACC 2612
OY 2457 CATATGCTGAAACCTCTTACTTTGTGTGTCAGATTTTGTGATTTTAAATTTCTACTTTC 2516
    |||||
```

```
Db 2613 CAGATCATCTGAACACTACTACTTTGTGTTCCAGA--TTTTTTTAAATCTCCTACTTC 2669
OY 2517 TCCTATCTTTGGGCAATCTGGGCACTTTTAAAGAGAGAAATGAGTGTGGTGAT 2576
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2670 TCCTATCTTTGAGCAATCTGGGCACTTTTAAATAAGAGAAATGAGTGAATGTGGTGAT 2729
OY 2577 AAACGTATATGTAAA 2591
    ||  ||||  ||||
Db 2730 CTGCTTTATCTAAA 2744
```

Search completed: March 19, 2002, 15:17:28
Job time: 4591 sec

P1 Darnell JE, Schindler CW, Shuai K, Wen Z, Zhong Z;
XX MPI: 1995-139598/18.
DR P-SDB: AAR72082.
XX Receptor recognition factor implicated in transcriptional
PT stimulation of genes - useful in drug screening assays and/or
PT for treating cellular debilitations, derangements and/or
PT dysfunction, etc.
XX
XX
PS Claim 10; Page 107-110; 160pp; English.
XX
XX A fragment encoding the human Stat91 protein was used to screen a
CC murine thymus and spleen cDNA for homologous proteins. A highly
CC homologous gene (given in Aa089338) was isolated that encoded a
CC 91 kDa protein (AAR72080) (Stat1) that was responsive to interferon-
CC gamma. Using a fragment of the mouse gene as probe, 2 additional
CC members of the 113-91 family of receptor recognition factor
CC proteins were isolated. The 2 genes (Aa089339-40) were cloned
CC in plasmids 13sfl and 19sfl and encoded proteins termed Stat4
CC (AAR72081) and Stat3 (AAR72082), respectively.
XX
SQ Sequence 2869 BP; 755 A; 743 C; 772 G; 599 T; 0 other;

Query Match 100.0%; Score 2869; DB 16; Length 2869;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2869; Conservative 0; Mismatches 0; Injels 0; Gaps 0;

QY 1 GCCGCGACGACGAGCCGCGCGCTCAGCGGCTCAGCCCGAGACAGTGGAGCCCTGACT 60
Db 1 gccgcagaccgacgagccgagccagtcgagctcagcccgagacagtcgagccctcagtc 60
QY 61 GCGAGGAGTGGCTCAGTGGAGAACAGTGCAGAGTGCAGACAGCTACTCTGAGAGAGC 120
Db 61 gccgcagagatgagctcagtcgagaccagctcagcagctgagcaacacgctactcgaagc 120
QY 121 TGCACGAGTGTAGAGGACGACGCTTCCCGATGAGCTGGGCGAGTTCCTGGACCTTGA 180
Db 121 tgcacgagctgtagcagagacgcttccccaagtcgagctgagcttccgacactgga 180
QY 181 TTGAGAGTCAAGAGTGGGCTATGAGAGCGAAGAGTGCACATGCGAGTGGTGTTC 240
Db 181 ttgagagtcgaagctggcagtcagcagcagcaagatcagcagctgagctgagctg 240
QY 241 ATATCTCTTGGGTGAATTTGACGACGATATACCGATTCTTGCAGAGTCCAGTCC 300
Db 241 ataactcttgggtgaatttgacacgaatataagcagatcttccgagagtlccaatgctc 300
QY 301 TCTATCGACACACTTTCGAAGATCAAGAGTTTCTGCAGAGAGTATCTTGAGAGC 360
Db 301 tctatcgacacacttccgaagatcaagcagtlctcgagagcagatccttgagaagc 360
QY 361 CAATGGAATTTGGCCGAGTGTGGCCGATGCTGTGGGAAGATCTCGCTCTCTCCAGA 420
Db 361 caatggaaatttgcccgagtcgtggccgatgctgtgggaagtlctcgctctccaga 420
QY 421 CCGCAGCCACGCGAGCCGACAGGGGGCCAGGCCAACACCAAGCCGCGCTATGGA 480
Db 421 ccgcagccacgagcccaagggggccagggccaaacccaacccaacgcccgtatga 480
QY 481 CAGAGAGCAGCAGATGTTGAGCAGCATCTTCAAGATGTCCGGAACGAGTGCAGATC 540
Db 481 cagagagcagcagatgttgagcagcatcttcaagatgtccggaagcgagtlcgagatc 540
QY 541 TAGAAGAGAAATGAGGTGGTGGAGAACTCCAGAGAGCTTTGATTTCACTACAAA 600
Db 541 tagaagagaaatgaggtgggtggagaaactccagagagcagacttgcattcaactaaaa 600
QY 601 CCCCAAGAGCCAGAGAGACATGCAGATCTGAATGGAACCAACAGCTGTGACCAAGC 660
Db 601 ccccaagagccagagagacatgcagatctgaatggaacaacacagctgtgacccagac 660

QY 661 AGAAGATGCGACGAGCTGGAACAGATGCTCAGACCCCTGGACGAGTGGAGAGCAATGG 720
Db 661 agaagatgcgacgagctggaaacagatgctcagacgcccctggacagatgaggaagcattg 720
QY 721 TGAGTGAAGTGGCGGGCTCTTGTCAAGCAATGAGTACGTGCAGAAAGACATCTGATG 780
Db 721 tgagtgaagtggcgggctcttgtcaagcaatgagtacgtgcagaaagacatctgatg 780
QY 781 AAGAGCTGGCTGACCTGGAAGAGCGGGCCAGAGATCCGTCATCGAGAGCCCTCCCAACA 840
Db 781 aagagctggctgacctggaaagagcgggcagagatcgctgcatcggaagcctcccaaca 840
QY 841 TCTGCTGAGACCTCTGGAAGAACTGATTAATTCATTAGCAGAAATCTCAATTCAGACCC 900
Db 841 tctgctgagacctctggaaagaaactgataatctcatagagatctcaactcagaccc 900
QY 901 GCCAACAAATTAAGAACTCGAGACCTCAGAGAAAGTGTCTCAAGCGGAGCTCA 960
Db 901 gccaaacaaatlaagaaactcgagagagctcgagagaaagtgctcacaagggtagccctca 960
QY 961 TCGTCGACGACCGCCCATGCTGAGAGAGAGATGCTGAGCTGTTCAGAACTTAATGA 1020
Db 961 tcgtcgacgacccgcccatgctgagagagatgctgagagatgctgagagacttaatga 1020
QY 1021 AGAGTGCCTTCTGTGTGAGCGGCGAGCCCTGCAATGCCATGACACCCGAGACCCCTTAA 1080
Db 1021 agagtgccttctgtgtgagcggcgagccctgcatgcccacgagacccgagccctaa 1080
QY 1081 TCAATCAAGATGAGTGTGACATTTACGCAAGTCAAGTGTGTCGTCGATTCCTGAGT 1140
Db 1081 tcaatcaagatgagtgtagcttaccacgaagtcaggtgctggtlccaactllccagagt 1140
QY 1141 TGAATTAACGCTTAAATTAAGGTGATGATAAAGCTCTGGGATGTTGCTGCC 1200
Db 1141 tgaattaacgctttaaattaaagtgatgatalaagaactctgggagatgttgcctgcc 1200
QY 1201 TCGAGAGTTCCTCGGAATTTAATCATCTGGGACGAAACAAAGTATGATACATGAGG 1260
Db 1201 tcgagagtctctcggaatttaattcatctgggacgaacaaagtagatgagacagag 1260
QY 1261 AGTGTAAACAGCGGAGCTGTCTGCAGAGTTCAAGACCTTGACCTTACGAGCAGAT 1320
Db 1261 agtgtaaacagcgagctgtctgcagagttcaagaccttgacctlaagagagagat 1320
QY 1321 GTGGGAATGAGAGCGCGTGCATTTGTATGCTTCCCTTGTATCTGACAGAGCTGCACC 1380
Db 1321 gtgggaatgagagcgcggtgcatttgtatgcttcccttgtatctgacagagctgcacc 1380
QY 1381 TGATCACCTTTCAGAGCTGAGGTGTACCAACAGGCTTCAAGATTGACCTACAGACCACT 1440
Db 1381 tgatcaccttctcagagctgaggtgtaccacaaaggcttcaagattgacctacagaccact 1440
QY 1441 CTTTGGCAGTTGTGTGATCTCCCAACTCTTTCAGATGCGCAAAATGCTTGGCATCATGCC 1500
Db 1441 ctttggcagttgtgtgatctcccaactcttctcagatgcgcaaaatgcttggcatcatgcc 1500
QY 1501 TGTGTATTAACATCTGACCAATTAACCCCAAGAGCTGAATCTTCTCACTAAGCCGCAA 1560
Db 1501 tgtgtatataacatctgacccaataaaccccaagagctgaacttctcaactaagccgcaa 1560
QY 1561 TTGGAACCTGGGACCAAGTGGCTGAGGTGCTCAGCTGGCAGTTCTGTCACACCAAGC 1620
Db 1561 ttggaacctgggaccaaagtggctgaggtgctcagctggcagttctgctcacaaccaagc 1620
QY 1621 GAGGCTGAGCATGAGCAGCTGCAACGCGTGGTGAAGAACTCTAGCGCTGATGGA 1680
Db 1621 gaggctgagcatgagcagctgcaacgctgggtgaagaaactctagcgctgatgga 1680
QY 1681 ACTACTCAGGCTGTACATCATGAGGCTAAATTTCTGCAAGAAACATGCTGCGCAAGC 1740
Db 1681 actactcagggtgtacatcatgaggttaatttctgcaagaaacatgctgcgcaag 1740
QY 1741 CCTTCTCTTCTGGGCTGCGCTACAGCAATATCATCGACCTTGTGAAAAGTATATCTTGG 1800

```
|||||
Db 1741 gcttcctctcgtggtcgtcgtacacatactcgaacctgtgaaaaaatacattcttg 1800
OY 1801 CCTTTGGAATGAAGGGTACATCATGGTTTTCATGACAAGAGCGGAGCGGCATCC 1860
Db 1801 cctcttggaatgaaggtacatcagtgttctacgcagaagcgcggtgacccatcc 1860
OY 1861 TAAGCAACAAAGCCCCGGGACCTTCTACTGCGCTTACGCGAGCAGCAAGAAAGAG 1920
Db 1861 taagcaacaagccccgggacaccttctactcgtcgtctcaagcgagagcaagaagag 1920
OY 1921 GGGTCACTTTTCATTGGGTGGAAAGACATCATGGCAAGACCCGATGTCATCTAG 1980
Db 1921 gggtaaccttctcacttggtggaagaagacatcagtgcaagaccgagatccagctcgt 1980
OY 1981 ACCCATACACCAAGCAGCAGCTGAACACATGTCATTTGGTAAATCATCATGGCTATA 2040
Db 1981 agccatacaaccaagcagcagctgaaacaacatgctcattgctgaatacatcagtggtcata 2040
OY 2041 AGATCATGATGCGACCAACATCTCTGTCTCCACTTGTCTACCTCTACCCGACATTC 2100
Db 2041 agatcatgatgacgaacacatcctgtgtctcacttgtctacacctacccgacatlc 2100
OY 2101 CCAAGAGAGAGCATTTGGAAGTACTGTAGGCCGAGAGCCGAGGACACCCGGAAGCG 2160
Db 2101 ccaagagagagcatctggaagaagtaactcgtcagccgagagccagagcacccgaaagcg 2160
OY 2161 ACCCAGATGATGTCGCCCGTACCTGAGACACAAAGTTCAATCTGTGACACCAAGACCT 2220
Db 2161 acccaggtagtgtgctgcccgtacctcgtgaagacaagttcatctgtcgtgacaccaagacct 2220
OY 2221 GCAGCAATACATTTGACCTGCGCATGTCCCGCGACCTTGTAGATTCATGTATGACAGTTG 2280
Db 2221 gcaagcaataccattgacctgcgcgaltgcccccgcacattagatcattgatgcagtttg 2280
OY 2281 GAAATACGAGTGAAGTGTGAGCCCTGACGAGAGGAGGAGTTGATGCTGCTACGTTG 2340
Db 2281 gaataacagtgtagagtgctcgaagccctcagcaagagcgagtttgagctcagcttg 2340
OY 2341 ACATGATCTGACCTCGAGGTGTCTACCTCCCGCATGTGAGAGAGCTGAAACGAGAAGCT 2400
Db 2341 acatgactcgtacctcgtgagtgctacctcccccatgtgagagagctgaacaacagaagct 2400
OY 2401 GCAGAGACCTGACTTGAACACCTGCGCGTGTCCACCCCTTAGACAGCCGAACCCCAT 2460
Db 2401 gcagagacgtgacttgagacacctgcccgtgtctcacaccttaagcagccgaaccccata 2460
OY 2461 TCGTCTGAACCTCTTACTTTGTGCTCCAGATTTTATTTTAAATTCCTACTTGTGCT 2520
Db 2461 tcgtctgaaacctctactcttggtgtccagatcttcttctaacttctcactctgct 2520
OY 2521 ATCTTTGGCAATCTTGCGACATTTTAAAAGAGAAATGAGTGTGCTGTGATTAAC 2580
Db 2521 atctttggcaatcttggtgaccttttaaaagagaaatgagtgagtggtggtgataaac 2580
OY 2581 TGTATGTAAAGAGAGAGACCTCTGAGTGTGGAGTGGGCGCTGAGACGAGAAGGAGCG 2640
Db 2581 tgtatagtaaagagagagacctctgagctcgtggtggtggtgagagagagagc 2640
OY 2641 AAAGGGGAACACCTCTCTGCTGCGCGCTGCGCTCTTTTTCAGAGAGTCCGGGGGTTGG 2700
Db 2641 aaaggggaacacctctcgtcctgcgcgcctgcccctcttctcagagagctcgggggttgg 2700
OY 2701 TTGTTTAGACAGTGCCTCTGCTGCTGATGCTTACCTGTTGCCCCATCTGTGAGCTGAT 2760
Db 2701 ttgttagaacaagtgcctcgtgctgccaatgtaacctgtgcccccaactcgtgagtgat 2760
OY 2761 ACCCATTTGGAAGTCTCTGCTGCTGACATTTGCAACCTTGCTAATATTCACACTTAAGC 2820
Db 2761 accccatctcgtggaacctcctcgtgctcgtcacttcaacctgtcctaataatacacaataagc 2820
OY 2821 TAGGACTAAGCCAGGAGGTTCTCTTAAATTAATAAAAAAAAAAAAAA 2869
|||||
```

```
Db 2821 taggactaagcccaaggaggtcctctctttaaattaaaaaaaaaaaaa 2869

RESULT 2
AAT31280
ID AAT31280 standard; cDNA; 2869 BP.
XX
AC AAT31280;
XX
AC AAT31280;
XX
DT 24-OCT-1996 (first entry)
XX
DE Mouse STAT4 cDNA clone 19sf6.
XX
KW STAT; STAT4; signal transducer and activator of transcription;
KW DNA binding protein; ligand; receptor; oncogenesis;
KW Inflammation; autoimmune disease; antagonist; therapy; ds.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 69..2381
FT /tag= a
XX
PN M09620954-A2.
XX
PD 11-JUL-1996.
XX
PF 28-DEC-1995; 95MO-US17025.
XX
PR 06-JAN-1995; 95US-0369796.
XX
PA (UYRQ ) UNIV ROCKEFELLER.
PI Darnell JE, Horvath CM, Wen Z, Zhong Z;
XX
DR MPI; 1996-33941/33.
DR P-PSDB; AAW03176.
XX
PT New STAT protein DNA-binding domain peptide(s) - useful for
PT diagnosing, preventing or treating cellular dysfunction, e.g.
PT oncogenesis, inflammation, parasitic disease or autoimmunity
XX
PS Disclosure: Page 93-96; 138pp; English.
XX
CC Mouse cDNA clone 19sf6 (AAT31280) codes for mouse signal transducer
CC and activator of transcription (STAT) protein STAT4 (AAW03180),
CC a protein that serves a dual purpose, i.e. signal transduction from
CC ligand activated receptor kinase complexes followed by nuclear
CC translocation and DNA binding to activate transcription. STAT4
CC includes a DNA-binding domain (see also AAW03167) useful for
CC screening for antagonists that modulate STAT activity.
CC The cDNA, which was obtd. from splenic/thymic cells, can be used
CC for prodn. of STAT4 or mutant STAT4.
XX
SQ Sequence 2869 BP; 755 A; 743 C; 772 G; 599 T; 0 other;

Query Match 100.0%; Score 2869; DB 17; Length 2869;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCGCGACCAAGCCAGCGCCGCGCTCAGCGCTCAGCCCGAGACAGTCGAGACCCCTGACT 60
Db 1 gccgcgaccagccagcgccgctcagtggtcagcccgagacagtcgagacccctgact 60
OY 61 GCAGCAGGATGGCTCAGTGAACACAGCTGACAGCTGGACACACGCTTACTTAAGCAGC 120
Db 61 gcagcaggatggctcagtgaacacagctgacagctggacacacgcttacttaagcagc 120
OY 121 TGCACACAGTGTACACGACAGCTTCCCATGTGAGAGTGGCGGACAGTTCCGTGACCTTGA 180
Db 121 tgcaacagctgtacacgacagcttcccatgtgagagtggtggcgacagttccgtgaccttga 180
OY 180
Db 180
```

0Y 181 TTGAGAGTCAGACTGGCATATGACAGCCAGAAAGTCCATGCGACGTTGGTGTTC 240
|||||
Db 181 tgaagatcaagactgagcatatgacagccagaaagatcaatgcaagcttggtgttc 240
0Y 241 ATAACTCTGGGTGAAATTGACCAATATATGCGGATTCTCGCAAGATCCAAATGCC 300
|||||
Db 241 ataactcttggtgaaattgaccacaatalatagccgattctcgcaagatccaaatgtcc 300
0Y 301 TCTATCAGCAACCTTCGAAGATCAAGCAGTTCTGCAGACAGATATCTTGAGAAC 360
|||||
Db 301 tctatcagcaaaccttcgaaagatcaagcagttctgcagagcaggtatcttgagaagc 360
0Y 361 CAATGGAAATATGCGCGGATGTCGCCGATGCTGTGGGAAGATCTCGCTCTCCAGA 420
|||||
Db 361 caatggaatatgcccggatgctgcccgatgctgtgggaagatctcgctctccaga 420
0Y 421 CGGACACCAAGCGAGCCAGCAAGGGGGCCAGGCCAACCCCAAGCGCGCGTAGTGA 480
|||||
Db 421 cggacaccaaagcagccagcaaggggccaagccaacccaacagccgctgtagtga 480
0Y 481 CAGAGAAAGCAGAGATGTTGAGACAGATCTTACAGATGTCGGAAAGCAGTCCAGATC 540
|||||
Db 481 cagaagaagcagcagatglttggagcaacatcttcaagatgtccggaagcgagtgcaagatc 540
0Y 541 TAGAAGAGAAATGAAAGTGTGTGAGAACCTCCAGAGCAGATTGATTTCACTACAAA 600
|||||
Db 541 tagaagaagaatatgaaagtgtgtgagaacctccagaagcagacttgatttcaactacaaa 600
0Y 601 CCTCAAGAGCCAGAGAGATCAGAGATCTGAATGGAACAAACAGTGTGACACAGAC 660
|||||
Db 601 cctcaagagccaaagagacatgcaagatctgaatggaacaacacagctgtgacacagac 660
0Y 661 AGAAGATGCAGCAGCTGGAAGAGGCGGCAAGATGCGGTGATCGGAGGCGGAGAGATG 720
|||||
Db 661 aagaagatgcagcagcctggaagagcgccagagacgctgagccagaatgcygaagagacatg 720
0Y 721 TCAGTAGCTGCGGCGGCTCTTGTTCAGCAATGAGTACGTGCGAAGACACTGATGATG 780
|||||
Db 721 tgaagtgcgcygagggctctgttgcagcaatgagtaagtcgcaagaacactgactgagt 780
0Y 781 AAGAGTGGTGTGCTGGAAGAGGCGGCAAGATGCGGTGATCGGAGGCGGCGGCAACA 840
|||||
Db 781 aagaagtcgctgcagcctggaagagcgccagagacgctgacgacgagagccctccaca 840
0Y 841 TCTGCTGACCGCTGTGAAAACCTGATTAATTCATTAGCAGAAATCTCAACTTCAGACC 900
|||||
Db 841 tctgctgacccgtctgaaaactgataactcattagcagaatctcaactcagaacc 900
0Y 901 GCCAAGAAATTAAGAACTGAGAGACTGCAGAGAAAGTGTCTTCAAGGCGGACCTTA 960
|||||
Db 901 gccaaagaaatlaagaaactggaagagctgcagcagaagaatgtctcctaagaaggcgacctla 960
0Y 961 TCGTGACAGACCGGCCCATGCTGTGAGAGAGAGATCGTGAGCTGTTTCAGAACTTAATGA 1020
|||||
Db 961 tctgtgacagacccggcccatgctgtgagagagagatctgtgagagctgttcagaacttaatga 1020
0Y 1021 AAGAGTCCCTTCTGTGTGAGAGCGGACCCCTGCATGCCATGCAACCCGGACCCCTTAG 1080
|||||
Db 1021 aagagtccctctgtgtgtgagcgcgacccctgcatagccccggacccgaccttag 1080
0Y 1081 TCAATCAAGACTGCTGTCCAGTTTACACAGCAAGTCAAGTTGCTGTGCAATTTCCCTGAGT 1140
|||||
Db 1081 tcaatcaagactgctgtccagtttacacagcaagtcaggttgcgtgcaaatctccgagtt 1140
0Y 1141 TGAATTAACAGCTTAATAATTAAGTGTGATTAAGACTCTGCGATGTTGTGTGCC 1200
|||||
Db 1141 tgaattatacagcttaataatlaaagtgtgatgataaagaaccttggtgtgtgtgcc 1200
0Y 1201 TCAGAGCGTCTGGGAAATTTAATTAATCTGCGACGCAACCAAAAGTGAATGAACATGAGAG 1260
|||||
Db 1201 tcagagcgctcgggaaatlaaataatctctgagcagacaaagaatgataatgaaagt 1260
0Y 1261 AGCTAACACGCGACGCTGTGTGAGAGTTCAGAGTTCAGACCTTGACCTTAGGACAGAGAT 1320
|||||

Db 1261 agcttaacaacgacagcctgtctgcagagttcaagcacctgaccccttaaggagcagagat 1320
0Y 1321 GTGGGAATGGAGGCGGTGCCAATTGTGATGCCCTTGTAGCTGAGTGGAGGCTCCACC 1380
|||||
Db 1321 gtgggaatggagcggtgcccaattgtgatgcccttgtagctgagtgagagctccacc 1380
0Y 1381 TGATCACTTCGAGACTGAGCTGTACCAACAGGCGCTCAAGATTGACCTAGAGACCCACT 1440
|||||
Db 1381 tgaactacttcgagactgaggtgtacccaagaagccttcaagattgacclagagaccact 1440
0Y 1441 CCTTGGCAGCTTGTGTGATCTCCAAATCTGTACAGATGCCAAATGCTTTGGCATCAATCC 1500
|||||
Db 1441 ccttggcagcttgtgtgtgattcccaacatctgtcagatgccaatgtcttggatccaatcc 1500
0Y 1501 TGTGTATTAACATGCTGACCAATTAACCCCAAGACCTGAATCTTCTCACTAAGCCGCCAA 1560
|||||
Db 1501 tgtgtataacatgtctgacccaataacccaagaagctgaaactcttcactaagccgca 1560
0Y 1561 TTGGAACTTGGACCAAGTGGCCGAGGTGCTCAGCTGGCAGTTCTGTCCACCACCAAC 1620
|||||
Db 1561 ttggaaccttggacccaagtgycgaagtgctcagctggcaggtctgtccacccaagc 1620
0Y 1621 GAGGCTGAGCATCGAGACGCTGACAAAGCGTGGCTGAGAACTCTAGGGGCTGTGTGA 1680
|||||
Db 1621 gaggtctgagcatcagagcagcttgacaacgctgtgagaagctccctgaagctgtgtga 1680
0Y 1681 ACTACTAGGSGTGCAGATCATGSGGCTAATTTGTGCAAGAAACATGCTGGCAAG 1740
|||||
Db 1681 actactcaaggtgtcagatcacatlggctaaatcttcgcaaaagaacacatgcttgcgaag 1740
0Y 1741 GCTTTCCTTCTGTGGTGTGGCTGACAAATATCATGCAACCTGTGAAAAATATATCTTGG 1800
|||||
Db 1741 gcttctccttcggtgtcgtgagacaataatcagaccttggaagaagatataatcttgg 1800
0Y 1801 CCTTTGGAATGAAGGATACATCATGCGTTTATCAGCAAGAGCGGAGCGGCCATCC 1860
|||||
Db 1801 cctttggaatgaaaggtacatcatggtttatcagcaagagcggtgcgtccatcc 1860
0Y 1861 TAAGCAAAAGCCCCGGCGACCTTCTACTGCGCTTTCAGCGAGACACCAAGAAAGAG 1920
|||||
Db 1861 taagcaaaagccccggcgaccttctactcgtcttcagcgagcgacaaagaagag 1920
0Y 1921 GGGTCACTTTCACCTTGGGTGGAAGAGACATCAGTGGCAAGACCCAGATCCAGTGTGTG 1980
|||||
Db 1921 gggctacttctactctgtgtggaagaagacatcagtgcaagaccagatccagctcgtlag 1980
0Y 1981 AGCCATACCAAGCAGCAGCTGACACATGTCATTTGCTGAATCATATGCGCTATA 2040
|||||
Db 1981 agccatacccaagcagcagcttgaacaacatgtcattctgaaatcatalcatgagctata 2040
0Y 2041 AGATCATGGATGGACCAACATCCTGTTGCTCAGTTGCTGTAACCTTCCCGGACATTC 2100
|||||
Db 2041 agatcatgatatgaccaaaccctgtgtctcactgtctcactcctacccctcagacatcc 2100
0Y 2101 CCAAGAGAGAGGATTTGAAAGTACTGTAGGCCCGAGACCGACGACCCCGAAGCG 2160
|||||
Db 2101 ccaagagagagagcatttggaaagtactgtlagagccgagagccagagcaccctcgagccg 2160
0Y 2161 ACCCAGGTAGTGTGCGCCCGTACTTAAGACCAAGTTTCACTGTGTGAGACCAAGAGCT 2220
|||||
Db 2161 acccaggtagtgtgccccgtactccttgaagaaccaagtcatctgttgaacacaaagacct 2220
0Y 2221 GCAGCAATACATGACCTGCGCATGTCCCGCGCACTTTAGATTTCATGATGAGTTTG 2280
|||||
Db 2221 gcagcaatcacatgtgacctgcgcgaatgtcccccgacacttgaatcatatgacgagtttg 2280
0Y 2281 GAAATACGCTGAAGGTGCTGAGCCTTCAGCAGGAGGAGGAGTTTGAAGTGCCTCAGTTTG 2340
|||||
Db 2281 gaataacggtgaaggtgtgtgagccctcagcaggaagcgagtttgaagtcgtccacglttg 2340
0Y 2341 ACATGATCTGACCTGTGGAGTGTGTACTCCCAATGAGAGACTGAAACCAACACT 2400
|||||

Db 601 cccctaaagaccagaacatcagagatctgaaatgaaacaacacgctctgagcacagc 660
Qy 661 AGAGATGACGACGCTGGAACAGATGCTCACAGCCCTGACAGATGCGGACAGCATTTG 720
Db 661 agaagatcgaagcagctcgtgacacagatcgtcacagccctgagccagatgacgagaagcatctg 720
Qy 721 TGAGTGACCTGGCGGGGCTCTTGTCAGCAATGAGTACGTGAGAAAGACACGACTGATG 780
Db 721 tgaatgaaatcgtgcgagcctctgtcaagcaatggaatcagtcgaagaacacagacatgagatg 780
Qy 781 AAGACCTGGCTACTGGAAGAGAGCGCCGACAGATCGCTGATCGAGCGCCCTCCCAACA 840
Db 781 aagagctgctgactcgtgaaagcgcgacagatcgctgcatcggaagccctcccaaca 840
Qy 841 TCTGCTGACCGCTCTGGAAAACTGATTAATTATTAGCAAAATTCATCTTCAACCC 900
Db 841 tctgctctgacgctcgtgaaaacttgataactcattagcaagaatcctaactcagacc 900
Qy 901 GCCAACAAATTGAAGAACTGAGAGAGCTCAGACGAAAGTGTCTCAAGAGCGGACCTTA 960
Db 901 gccacaacaatgaagaacatcgagagagctcgacagagaagtgctccatacaagcgaccccta 960
Qy 961 TCGTGACGACCGCGCCCATGCTGAGAGAGAGATCGTGGAGCTGTTACAGAACTTAATGA 1020
Db 961 tctgtcagacacgcccacatgctgagagagatcgatgagctgttcagaaaactaatga 1020
Qy 1021 AAGATGCTTCTGATGATGAGAGCGGACCGCTGATGCCCATGACACCGGACCGCCCTTAC 1080
Db 1021 aagatgctctctgtgtgagcgagcgagcctcgtacccatgcaacccgagacgacctag 1080
Qy 1081 TCATCAAGACTGCTGTCTCAGTTTACACGAAAGTCAAGTGTCTGCTCAAAATTTCTGACT 1140
Db 1081 tcatcaagactgctgtctcagtttaccacgaaatcagtgctgtgtcaaaattctcagat 1140
Qy 1141 TGAATATATAGCTTTAAATTTAAAGTGTGATGATTAAGCTCTGGGAGTGTGCTGCTCC 1200
Db 1141 tgaatataagctttaaatttaaagtgatgataaagctctgggagatgtgtgcgccc 1200
Qy 1201 TCGAGAGGCTCTCGGAAATTTAACATTTCTGGGACGAAACACAAAAGTATGATGAGAG 1260
Db 1201 tcgagagctctcgaaatttaacatctctgggacgaaacacaagaatgatgaatgagag 1260
Qy 1261 AGCTTAACAACGCGACGCTCTCTGACAGATTCAAGCACTGACCTTAAGGACAGAGAT 1320
Db 1261 agcttaacaacgcgacgctctctgacagagttcaagcacctgaccccttaagagacagat 1320
Qy 1321 GTGAGATGAGAGCGCGGCAATTTGATGCTCTGCTGATGCTGACTGAGAGAGCTGCACC 1380
Db 1321 gtgagatatgagagcgctgtccaattgtatgctccctctgactcgtgactgagagctgaccc 1380
Qy 1381 TGATCACTTCGAGACTGAGGTGTACCAACAGGCTCAAGATTGACCTAGAGACCCTACT 1440
Db 1381 tgaatcacttcgagactgaggtgttaccaacaagcctcaagatgtacatagagaccact 1440
Qy 1441 CCTTGGCAGTTGGTGATGATGTCACATCTGTCAGATGCCAAATGCTTGGGCAATCAATTC 1500
Db 1441 ccttggcagttggatgataatgtaacatctgtaagaatgccaatgcttgggcatcaatcc 1500
Qy 1501 TGTGATATAACATGCTGACCAATTAACCCAGAAAGCTGAATCTTTCTACATAAGCGCCCAA 1560
Db 1501 tgtgataataacatgctgaccaataaccccaagaagcgtgaactcttctacataagcgcgcaa 1560
Qy 1561 TTGGAACCTGGGACCAAGTGTGGCGGAGGTGCTCAGCTGGCAATTTCTCTCCACACCAAGC 1620
Db 1561 ttggaacctgggaccaaagtgtggcgaggtgtctcagctgtgacagctctgtccacaacaagc 1620
Qy 1621 GAGGGCTGACATCGAGCAGCTGACAAAGCTGCTGAGAAAGCTCTAGGGCGCTGTGTGA 1680
Db 1621 gagggctgacatcgagcagctgacaaagcgtgctgagaaagctctcctagggcctgtgtga 1680
Qy 1681 ACTACTGAGGGTGTGAGATACATGGCTAAATTTCTCAAGAAAGAAACATGCTGCTGGCAGG 1740
Db 1681 actactgaggggtgtgagatcacatgagctgaaattctgcaagaagaaacatgctgtgcaag 1740

Db 1681 actactcaaggtgtcagatcacatgagctgaaattctgcaagaagaaacatgctgtgcaag 1740
Qy 1741 GCTTCTCCTTCTGGGCTGTGGTACCAATATCATGACCTTGTGAAAAGTATATCTTG 1800
Db 1741 gcttctccttctgggctgtggatccaataatacatcagacctgtgtgaaaagtataatccttgg 1800
Qy 1801 CCTTTGGAATGAAGCGTACATCATGCGTTTTCATCAGCAAGACGCGGAGCGGCTATCC 1860
Db 1801 cctttggaatgaaaggtatcatcatatggttcatcagaagaagcgagagcgagccatcc 1860
Qy 1861 TAAGCAAAAGCCCCCGGACCTTCTACTGCGCTTTCAGCGGAGACGACCAAGAAAGAGAG 1920
Db 1861 taagcaaaaagccccgggacaccttctactcgtgcttcaagcgagagcagaagaagagag 1920
Qy 1921 GGGTCACTTTCATCTGGTGGGAAAAAGACATCAGTGGCAAGAACCCAGATCCACTGTGAT 1980
Db 1921 gggctacttctacttgggtgggaaaaagacatcagtgagcaagaccagatccagctctgag 1980
Qy 1981 AGCCATACACCAAGCAGCAGCTGACACATGCTCATTTGTGAATATCATATGCGCTATA 2040
Db 1981 agccatacaccaagcagcagctgaaacaatgtcatctgtgaataatcatatgagctata 2040
Qy 2041 AGATCATGATGCGACCAACATCTGTGCTCAGCTTGTCTACTCTACCCGACATTC 2100
Db 2041 agatcatgatatgacccaacatccctgtgtctcactgtgtcactctacacctacccgacatcc 2100
Qy 2101 CCAAGAGAGAGGCAATTTGGAAGTACTGATAGCGCCGAGAGCCAGAGACCCCGAAGCG 2160
Db 2101 ccaagagagagcatcttgaaagtactgtagccgagagccagagcaccgccgaagccg 2160
Qy 2161 ACCCAGTACTGCTGCCCGCTACTTGAAGACCAAGTTCACTGCTGTGACACCAACACCT 2220
Db 2161 acccagtgatgctgtcccgctacttgaagacccaagttcaatctgtgtgaaaccaagaccl 2220
Qy 2221 GCAGCAATACCATTTGACCTCCGATGTCCCGCCGCACTTGTAGATTGATGATGACCTTGG 2280
Db 2221 gcagcaataacatctgacctgtgcgagtgtcccccgcactttagatctatgtatgacgttgg 2280
Qy 2281 GAAATAACGCTGAAGGCTGTGAGCCTCTCAGCAGAGAGGCGATTTGAGCTGCTACGTTTG 2340
Db 2281 gaataaacgctgaaggctgtgcgaagccctcagcaagagagcagcttgatgtgtcctacgttgg 2340
Qy 2341 ACATGATCTGACCTCGAGCTGTGCTACCTCCGCCATGTAGAGAGCTGAAGCAGAAAGCT 2400
Db 2341 acatgatctgacctcgagctgtgctacctcccccacatgtgagagagctgaaaccaaagact 2400
Qy 2401 GCAGAGAGCTGACTTGAGACACCTGCCCGCTGCCATCCCTTAAGCAGCGCAACCCCAATA 2460
Db 2401 gcagagagctgacttgagacacctgtgcccggtgtccacccctlaagcagccgaaccccata 2460
Qy 2461 TCGTGTAAACTCCTACTTTGTGTGTTCCAGATTTTCTTTTAAATTTCTACTTGTGCT 2520
Db 2461 tctgtgtaaactcctactttgtgtgttccagatcttcttcttcttcttcttcttcttcttct 2520
Qy 2521 ATCTTTGGCAATCTGGGCACTTTTAAAGAGAGAAATGATGATGCTGTGATTAAC 2580
Db 2521 atctttgggcaactctggaccttctttaaagaagagaaatgagtgagtggtgtgtatgaac 2580
Qy 2581 TGTATGTAAAGAGAGAGACCTCTGAGTGTGGGATGAGGCTGAGAGCAAGAGGAGGC 2640
Db 2581 tgtatgttaaagagagagacctctgagcttggagatcgtggagcgtgagagcagaagggagagc 2640
Qy 2641 AAAGGGAACACGCTCTGCTGCGCGGCTGCCCTCTCTTTTTCAGAGCTCGGGGGGTGG 2700
Db 2641 aaagggaacacgctctgtctgcgcgcgctgcctctcttctcaagagctcgggggtgtgg 2700
Qy 2701 TTGTTAGACAAGTGCCTGCTGTGTGCCATGCTACTGTTTGCCCACTGTGTAGCTGAT 2760
Db 2701 ttgttagacaagtgctctccgtgtgtcccatgtctacacgtgtgcccacatctgtgtagctgat 2760
Qy 2761 ACCGCATTTCTGGGAACCTGCTGCTGTGCACTTTCAACTGTGTAATTCACCAATPAGAGC 2820
Db 2761 accccatctcggaacatccgtgctgtgcaacttcaacactgtctaataccacataagaagc 2820

[illegible]

OY	131	GTACACGCTTCCCATGAGACTGGCGGACGTTCCGAGACCTTGATTTGAGAGCA	190
Db	321	gtacagagcaagcttccccatgagctcgcgcgaagcttcctgcaccttgatctgaagca	380
OY	191	AGACTGGCATATGACAGCCAGCAAAAGATACATGTCACGTTGGTGTTCATATCTTT	250
Db	381	agactggcatatgacagccagcaaaagatcacatgacacgcttggtgttcatatcctt	440
OY	251	GGGTGAATAATGACCAACCATATATGCGGATTTCTTCGAAGAGTCCATGTCTCTATCACA	310
Db	441	gggtgaataatgaccacacatattatgcccgaatttcttcgaagagtcctatcaca	500
OY	311	CAACCTTCGAAGATCAAGCATTTCTTCGACACACGTTCTTGAAGAACCATGCAAT	370
Db	501	caaccttcgaagatcaagcatatttcttcgaagagatctcttgaaagccaatgcaat	560
OY	371	TGCCCGGATCGTGCGCCGATCGCTGTGGGAAGTGTGCGCTCTCTCAGACGGCAGCCAC	430
Db	561	tgcccgatctgtgcgcgaatgctgttggaagatctcgcctctccagaagcgaccac	620
OY	431	GGCAGCCACGACAGGGGGCCAGGCCAACACCCACAGCGCGGTAGTACAGACAACA	490
Db	621	ggcagccacgacagggggccagggccaaacccacacagctgcgtatgacagaaaca	680
OY	491	GCACATGTTGGAGCAGCATTTCAGATGTCCGGAAGCCAGATGCAAGATCTGAACAGAA	550
Db	681	gcacatgttgagcagcatttcagatgtccggaagccagatgcaagatctgaacaga	740
OY	551	AATCAAGTGGTGGACACACTCCAGCACGACTTTGATTTCACATAAACCTTCAGAG	610
Db	741	aatcaagtgttgacacactccagcacgactttgattttcacataaaaccttcagag	800
OY	611	CCAACATCAACATGCGAGATGTGATGTGAAGAAACACCGATGTGGACACAGACATGTCA	670
Db	801	ccaacgaaacatgcagatctgaaatggaacaaacccgctctgacacagaaagatgca	860
OY	671	GCAGACTCAACGATGCTCAACAGCCCTTGACACGATGCGAGAAAGCATTTGATGATGACT	730
Db	861	gcagactcaacgatgctcaacagcccttgacacgatgcgagaagcatgttgatgact	920
OY	731	GGCGGGGCTCTTGTCACCATGGAGTATACGTGCGAAGACACTGACATGATGAAGACTGCG	790
Db	921	ggcggggctcttgtcaccatggagtatacgtgcgaagacactgacatgaaagactgcg	980
OY	791	TGACTGGAAGAGCGCGCAGAGATCGCGTGTCATCGAGAGCCCTCCCAACATCTGCTGGA	850
Db	981	tgactggaagagcgcgagagatcgcgtgcatcgaagcgcttccaacatctgcctga	1040
OY	851	CCGCTCTGGAAACCTGGAATCTTACTTATAGAGAACTCAACTTCAGACCCGCCAACAAAT	910
Db	1041	ccgctctggaacctggaaacttacttattagaagaatctcaacttcagaaacccgaacaat	1100
OY	911	TAAACAACCTGGAGAGCTGCGACGCAAAAGTGTCTCAAGAGGCGACCTATGTGCACAGA	970
Db	1101	taagaacacttgagagagcttcgacgaagaagtctctacaagagggaacctatcgtgcaga	1160
OY	971	CCGCGCCATGCTGGAGAGAGAGATCGTGAGCGCTTTTCAGAAACTTAATGAGAGTGCCTT	1030
Db	1161	ccgcgcctatgcttgaagagagatcgtgagagctgttccagaacttaatbaaagatgctt	1220
OY	1031	CGTGGTGGAGCGGAGCGCTTGCAATGCCATATGACCGCGAGCCGGCCCTTGTATCAACAC	1090
Db	1221	cgtggtagagcgagccgcttgcaatgcccatatgaccgagccggccctctagatcaaaac	1280
OY	1091	TGCTGTCCAGTTTACACAGAAAGTCAAGTGTGCTGTCAAATTTCCTGAGTTGAATTATCA	1150
Db	1281	tgctgtccagtttacacagaaagtcaagtgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgt	1340
OY	1151	GCTTAAATTTAAAGTGTGCATGTATAAAGCTCTGGGAGATGTTGCTGCGCTTCAGAGGCT	1210
Db	1341	gcttaaatttaagtgtgcatgtataaagctcttgcgtgtgtgtgtgtgtgtgtgtgtgtgtgt	1400

[illegible]

Dd	2481	tgaagtgctgtagccctcaagcaggagggcaggtttagtcgcctcacgattgacatgatcat	2540
Oy	2351	GACCTGGAGTGTGCTACTCCTCCCATGTGAGGAGCTGAACACAGACTGCAGAGACT	2410
Dd	2541	gacctcgagtgtagctactcctcccattgtgaggagtgaaccagaagctgcagagacgt	2600
Oy	2411	GACTTGAGACACACTGCCCGTGCTCCACCCTCAACGACCGCAACCCCAATC	2462
Dd	2601	gacttgaagacacctgccccgtgctccaccctlaagcagccgaacccaatalac	2652
<hr/>			
RESULT 5 AAT05616 ID AAT05616 s tandard; CDNA to mRNA; 2787 BP.			
XX	AA05616;		
XX	AAT05616;		
DT	01-APR-1996	(first entry)	
XX			
DE	Human placenta acute phase response factor gene.		
XX			
KW	Human; acute phase response factor; transcription factor; interleukin-6; signal transmission; placenta; antibody; antisense; ribozyme; antiinflammatory; antitumour; hypotensive therapy; ss.		
XX			
OS	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	CDS	221..253	
FT		/tag= a	
FT		/product= Acute phase response factor	
FT		/note= "Open reading frame (claim 7) "	
XX			
PN	EP676469-A2.		
XX			
PD	11-OCT-1995.		
XX			
PF	29-MAR-1995;	95EP-0104670.	
XX			
PR	04-APR-1994;	94JP-0065825.	
XX			
PA	(KISH/) KISHIMOTO T.		
PI	Akira S, Kishimoto T;		
DR	WPI: 1995-346089/45.		
DR	P-PSDB; AAR82993.		
PT	New acute phase response factor - for developing inhibitory agents		
PT	for treating diseases induced by cytokine(s) such as IL-6, e.g.		
PT	inflammatory diseases		
XX			
PPS	Claim 8; Page 14-15; 31pp; English.		
XX			
CC	The sequence encodes a human acute phase response factor (APRF), a		
CC	transcription factor related to signal transmission of Interleukin-6		
CC	(IL-6). Human placenta cDNA is isolated using an IL-6-treated mouse		
CC	liver polymerase chain reaction product as a cDNA probe. APRF-		
CC	inhibitors, e.g. antibodies, antisense oligonucleotides or ribozymes,		
CC	may be used to treat diseases induced by IL-6, e.g. inflammatory		
CC	disease, leukemia, cancer, osteoclastia, pulmonary hypertension, etc.		
XX			
SQ	Sequence 2787 BP; 729 A; 719 C; 753 G; 586 T; 0 other:		
<hr/>			
Query Match 74.7%; Score 2144.2; DB 16; Length 2787;			
Best Local Similarity 90.9%; Pred. No. 0;			
Matches 2205; Conservative 0; Mismatches 223; Indels 7; Gaps 2;			
Oy	61	GCAGCAGATGGCTCAGTGGACACGCTGCAGCAGCTGCACACACGCTTAAGTCGAAGCACG	120
Dd	213	gaaacaagatggtgcccaatgaaatcagctacacagcagctgacacacgtaacctgagagcagc	272

[illegible]

QY	1201	TCAGAGGCTCTCGGAATTTTACATTTCGTGGCGACGACAAACAAATGATGTAACATGAGG	1260
QY	1201	TCAGAGGCTCTCGGAATTTTACATTTCGTGGCGACGACAAACAAATGATGTAACATGAGG	1260
Db	1353	lcaaggatcccggaataatlttaacattcttgygacaacaacaagaatgtatgaacttggag	1412
QY	1261	AGTGTAAACAAGGAGGCGCTTGTGTGCAGAGTTCGAAGCACCAGTGCCTTAGGACGAGAGAT	1320
Db	1413	aatccaacaacgagcctctctgcgaatltcaaacacttgcaccttggaggagcagat	1472
QY	1321	GTGGGAATGAGAGCGCTGTGCCAATTGTGATGCTCTTGATCTGTGACTGAGAGCTGCACC	1380
Db	1473	gtggaaatbgyggcgccagccaatttgaatgtcttcccttgaatttgaacttgaagagctgcacc	1532
QY	1381	TGATACACCTTGCAGACTGAGGTGTACCAACCAAGGCGCTCAAGATTGACTAGAGGCCACT	1440
Db	1533	tgaacccttgaagaccggaagtgtatcaccaaggtctccaagatltgaccuagagaccact	1592
QY	1441	CGTTGCAGTTGTGTGATCTCCAGACATCTGTTCAGATGCCAATGCTTTGGGCATCAATCC	1500
Db	1593	cctltgacgttltgtgtatctccaacaatctgtccaagatgtccaaatgtccttggcgltccatcc	1652
QY	1501	TGTGGTAAATCATGCTTCAGCAATTAACCCCAAGAGCTGAACTTTCACCTTAAGCGGCAA	1560
Db	1653	tgtgttaacaacaatgtgtgacccaacaacccaagaatgtgaacttcttccactaagcgcgcaa	1712
QY	1561	TTTGAACCTTGAGGACCAAGTGGCCGAGGTGCTCAGCTGGAGGAGTTTCGTCCACACCAAGC	1620
Db	1713	tttgaaccttgygaccaagtggtccggaagtgtcagctgtgagttctgtcttccaccacaagc	1772
QY	1621	GAGGGCTGTAGCATATGAGACACTACACAACCTGTGGTGTAGAAGCTCTTAGAGGCTGGTGTGA	1680
Db	1773	gagggctgtagcatatgagacactacacaacctgtgggtgtagaagctcttagaggctgggtgtga	1832
QY	1681	ACTACTCGGCTGTCACATATACATGTGGGTAAATTTGCAAAACAAACATGTGCTGGCAAG	1740
Db	1833	actactcgggtgttccaatatacatatctgttcaactcttgcagaagaataactatgtgttgcgaag	1892
QY	1741	GCTTCTCCTTCTGGGTTGTGCTTAGACAAATATCATGACCTTGTGAAAAGTATATCTTGG	1800
Db	1893	gcttctccttacttgygtctgtagacaataatcatctgaccttltgaaaaagtatactlttg	1952
QY	1801	CCCTTTGGAATGAAGGGGTATCATGTGGTTCATATCAGCAAGAGGGGAGCGGGGCATATCC	1860
Db	1953	cccttltggaatgaagggtatcatgtgggttcattatcagcaagaggagcgaggagcattcl	2012
QY	1861	TAAACCAAAAGCCCCCGGGACCTTCTTACTGTGCGCTTAGCGAGAGCAGCAAAAGAGAG	1920
Db	2013	tgaagcaaaagcccccgaggacaccttctctgtggtcttcaagttaaagcagcaagaagaagag	2072
QY	1921	GGGTACCTTTACATTGGGTGTGAAAAAGCATACATGTGTGGCAAGCCGAGATCTGTAG	1980
Db	2073	gggtacctttaacttgygttgaagaagaatcagtgagtaagaccagatccaagtcgcgtg	2132
QY	1981	AGCCATACACCAACACACACACTGAACAACATGTCAATTTGCTGAATTCATCATGGGCTATA	2040
Db	2133	aacctatacaaaagcagcagctgacaaataatgttcalcttgcigaataatcatatgtggtcata	2192
QY	2041	AGATCATGTGATGCGACCAACATCTCTGTGTCTTCACATTGTCTACTCTTACCCGCACATTC	2100
Db	2193	agatcatgtgatgtcgacaataatctctgttcttccactgttcatctatcatcttgaatc	2252
QY	2101	CCAAGGAGGAGGCAATTTTGGAAAGTACGTGTAGAGGCCGAGAGCCAGAGAGACCCCGAAGCG	2160
Db	2253	ccaaggaggaggaacttcgggaaatatttgcgcgcgaagagcgcaagagatccctgtaaagctg	2312
QY	2161	ACCCAGTACTGTGCCCGGTACTGTAAGACCAAGTTCACTGTGTGTGACACCAAGACT	2220
Db	2313	acccagtgagtgctgcctccatactcgtgaagacaagttatctgttgacaccaagacct	2372
QY	2221	GCACGAATACCATTTGACCTTCGCATGTCCCCCGCAGCTTTAGATTCATTTGACAGTTTG	2280
Db	2373	gcacgaataatcatltgaccttgcgaatgtcccccgcgtttagatcatcttgaatgtgaatttg	2432
QY	2281	GAATTAACGGTGAAGGTGTGACGCTCTTCAGACAGAGAGGAGAGTTTGAAGTCGCTACGTTTG	2340

Db 2433 gaaataatgtaaggtgctgaacccctcagcaggaagcaglttgctccctaccttg 2492
Oy 2341 ACATGATGCTGACCTGAGGTGTGTAACCTCCCGCATGTGAGAGGTGAACCGAAGCT 2400
Db 2493 acatgagtgtagctgaggtgctgctacccctcccatgtgaggaagctgagaagagct 2552
Oy 2401 GC---AGAGACGTGACTTGAGACACCTGCGCCGTGCTCCACCCCTTAAGACGCCGAACC 2456
Db 2553 gcgaaagatacgaactgagcgcctacatcgtcctcgcacccctccacacagccaaacc 2612
Oy 2457 CATATGCTGGAACCTGCTTAACCTTTGGTTCAGATTTTCTTTTAAATTCTTACTTC 2516
Db 2613 caatacctcgaacactaactctgtgtccaga---tttttctaactcctacttc 2669
Oy 2517 TGCTATCTTTGGGCAATCTGGGCACTTTTAAAGAGAGAATGAGTGTGCTGGTGTAT 2576
Db 2670 tgcatactttagaacaatctgagcactttaaataagaagaatgagtgtgtgtat 2729
Oy 2577 AACCTGTTATGTA 2591
Db 2730 ctgcttatacctaa 2744

RESULT 6
AAF21060
ID AAF21060 standard; DNA: 2787 BP.
AC AAF21060:
XX
DT 14-MAR-2001 (first entry)
XX
DE Human low adenosine antisense oligonucleotide related sequence #2627.
XX

KM Low adenosine antisense oligonucleotide: phosphorothioate; allergy;
XX human; airway disorder; bronchoconstriction; lung inflammation;
XX surfactant depletion; respiratory; bronchodilation; antiinflammatory;
KM immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
KM respiratory obstruction; pulmonary obstruction; impeded respiration;
KM surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KM respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KM chronic hypertension; emphysema; pulmonary transplantation rejection;
KM chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KM cancer; SS.
XX
OS Homo sapiens.
XX
PN WO200062736-A2.
XX
PD 26-OCT-2000.
XX
PF 24-MAR-2000; 2000WO-US08020.
XX
PR 06-APR-1999; 99US-0127958.
XX
PA (UYEC-) UNIV EAST CAROLINA.
XX (NYCE/) NYCE J W.
XX
PI NYCE JW;
XX
DR WPI: 2000-679539/66.
XX
PT Low adenosine (A) content antisense oligonucleotides which do not
XX trigger adenosine receptors during metabolism, useful e.g. for treating
PT cancers and respiratory obstructions -
XX
PS
XX Disclosure; Page 867; 1592pp; English.

XX The present invention describes low adenosine (A) content antisense
CC oligonucleotides and compositions (1) comprising them, in the antisense
CC oligonucleotides the A is replaced by a 'universal' or alternative base.
CC (1) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.

CC The antisense oligonucleotides and (1) can be used to down-regulate the
CC expression and/or activity of target polypeptides associated with
CC lung/respiratory disorders and malignancies, such as stimulating and
CC activating peptide factors and transmitters, transcription factors,
CC immunoglobulins and antibodies, antibody receptors, cytokines and
CC chemokines, endogenously produced specific and non-specific enzymes,
CC binding proteins, adhesion molecules and their receptors, cytokine and
CC chemokine receptors, adenosine receptors, bradykinin receptors, central
CC nervous system (CNS) and peripheral nervous and non-nervous system
CC receptors, CNS and peripheral nervous and non-nervous system peptide
CC transmitters, defensins, growth factors, vasactive peptides and
CC receptors, binding proteins and malignancy associated proteins. The
CC antisense oligonucleotides may be used in this way to treat disorders
CC including respiratory obstruction (especially pulmonary obstruction
CC and/or bronchoconstriction) and/or lung inflammation, allergies)
CC and/or surfactant hypoproduction which are associated with a disease or
CC condition selected from pulmonary vasoconstriction, inflammation,
CC allergies, asthma, impeded respiration, respiratory distress syndrome
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC fragments and antisense oligonucleotides used in the exemplification of
CC the present invention.
XX
SQ Sequence 2787 BP; 729 A; 719 C; 753 G; 586 T; 0 other:

Query Match 74.7%; Score 2144.2; DB 21; Length 2787;
Best Local Similarity 90.9%; Pred. No. 0;
Matches 2305; Conservative 0; Mismatches 223; Indels 7; Gaps 2;

Oy 61 GCACGAGATGCTCACTGCAACCACTGACAGCTGGACACAGCTTACTGACAGC 120
Db 213 gaaacagatgagcccaatgagatcagctacagcagcttgacacagctgagagc 272
Oy 121 TGACACAGCTTACAGGACACGCTTCCCATGAGCTGGCGGACGCTTCCGACCTTGA 180
Db 273 tccatcagctcctacagagcagctcccaatgagagcagcttctcgccttgga 332
Oy 181 TTGAGAGTCAAGACTGGGCAATATGACAGCCAGCAAGATGCATGCAGTGTGTTTC 240
Db 333 ttgagagtaagaattgagcatalgagcagcaagaatcacatgcacttggtgttc 392
Oy 241 ATATCTCTTGCGTGAATTTACAGCAATATACCGCATTCCTGCAAGAGTCAATGCC 300
Db 393 ataattcctcggagagatgagcagcagatagcgccttcgcaagagtcgaatg 452
Oy 301 TCTATCAGACACACCTTTCAGAGATCAAGAGTTCCTGACAGCAGGTATCTGAGAAC 360
Db 453 tctatcgcacaatcctgaagaatcaagaagcttctcagagcagatcttgagaagc 512
Oy 361 CAATGGAATTTGCCCGATGCTGCGCGATGCTGTGGGAAGAGTCTGCTCTCCAGA 420
Db 513 caatgagatgcccggatgctgcccggatgctgctggaagaatcacagccttcaaga 572
Oy 421 CGGACGCCAGCGCCAGCAGAGGGGCGAGGCCCAACACCAACAGCCCGCTAGTGA 480
Db 573 ctgacgcaatcgggccagcagaagggcgccagccacaccccaagcagcagctg 632
Oy 481 CAGAGAACGACAGATGTTGAGCAGCATCTTCAGAGATGTCGGAAGCAGTGCAGCATC 540
Db 633 cggagaaagcagcagatgctgagcagccttcaagatgctcggaaagagctgagc 692
Oy 541 TAGAACAAGAAATGAGTGTGGGAGACGTCACAGCAGCTTGTGATTTCAATTCACAAA 600
Db 693 tagaacagaaaatgaaatgagatgagaaatcccaagatgacttgaattcaactataaa 752
Oy 601 CCTCAAGAGCCAGGAGACATGCAAGATCTGAATGAGAAACCAAGCTGTGACACAGC 660
Db 753 cctcaagaagcagagacatgcaagatctgaaatggaacaacacagcagtgaccagc 812
Oy 661 AGAAGATCAGCAGCTGGAACAGATGCTCACAGCCCTTGACCAAGATCGGGAAGCATTTG 720

|||||
Db 813 agaagatgcagcagctcgtgaacagatgctcactgcctgagccaatgcggaagcaccg 872
OY 721 TGAGTGAAGTGGCGGGCTCTTGCAGCAATGGAGTACGTACGAGAAAGCACTGACTGATG 780
Db 873 tgaatgagctcggcgggctcttctgcaagcagatggagtaagtcagaaaactctcaagcag 932
OY 781 AAGAGCTGGCTGACTGGAGAGCGCGCCAGAGATCGCTGCATCGGAGCGCCCTCCCAACA 840
Db 933 agggagctgcgtactgcgaagagcgcgcaacagatctgcctgcatlggaagcccgcccaaca 992
OY 841 TCTGCTTGAGACCGTCTGGAAGAACTGGATTAATTGATGAGCAATCTCACTTCAGACCC 900
Db 993 tctgcttagatcgcgtagaacacgcgaaacgtaacgtaactagcagaactcctaactcaagacc 1052
OY 901 GCCAACAAATTAAGAAACTGAGAGAGCTGCAGCAAGAAAGTGTCTACAAAGCGACCTTA 960
Db 1053 gtcaacaataatgaagaactggaagagtgatgcaccaaaaagttctctacaagaaggagacccca 1112
OY 961 TCGTGCAGCACCGGCCCATGCTGAGAGAGAGATCGTGGAGCTGTTCAAGAACTTAATGA 1020
Db 1113 ttgtaacagcacccgagccgagtcgtgagagagatcgtggaagctgtctcagaacttaatga 1172
OY 1021 AGATGCTCTTGCTGTGGAGGGGAGCCCTGCTATGCCATGACCCGAGCGCCCTTAG 1080
Db 1173 aaagtcctcttggtgagcgagcccgatgcatacctgcatacctgcagccgcccctcg 1232
OY 1081 TCATCAAGACTGCTGTGCTGAGTTTACACAGAAAGTCAGTGTCTGGTCAATTTCTGAGT 1140
Db 1233 tcatcaagaacggcgctcagctcactactactaaagtcagtgctgctcagatctccctgagt 1292
OY 1141 TGAATTTATCAGCTTAAATTAAGTGTGATTAAGTAAAGCTGTGGAGTTCCTGCGCC 1200
Db 1293 tgaattatcagcttaaaaatlaaagtgctgcatlgacaagaactcgggaagctgtgagctc 1352
OY 1201 TCAGAGGCTCTCGGAAATTTAACTTCTGGGACAGAAACACAAAAGTGAATGAACATGAGG 1260
Db 1353 tcaagagatcccggaataatlaacatctcggcacaacaacaaagtgatgaacatggagag 1412
OY 1261 AGTGTAAACAGCGAGCTGTCTGTGAGATTCAAGCACCTGACCTTGGGAGCGAGAGAT 1320
Db 1413 aatcaaaaaagcgagcctcctcgtcagaagatccaactctgaccctgagggagcaagat 1472
OY 1321 GTGGGATGAGAGCCCGTGCATTTGTATGCTCTTGATCGTGAATGAGAGCTGCACC 1380
Db 1473 gtgggaatggggcgagccaatctgtagctctccctgtagctgtagaagagctgcgacc 1532
OY 1381 TGATCACCCTTGAGAGCTGAGGTGATACGACCAAGGCTCAAGATGACCTAGAGCCCACT 1440
Db 1533 tgaatcaactcttgagaacgagtgatcaccaaggtctcaaatctgacctagaagaccact 1592
OY 1441 CCTTGCCAGTTGTGTGATCTCCAAACATCTGTAGATGCCAAATGCTTGGCATCAATCC 1500
Db 1593 cctgtcagtggtgtagtctccaacatctgtcagatgcacaatgacctggcgctccatcc 1652
OY 1501 TGTGTATTAACATGCTGACCAATTAACCCAAAGAGTGAACCTTCTTCACTAAGCCGCAA 1560
Db 1653 tgtgtatacaaatctgtacccaacaatcccaagaatgtgaactctctcactaagcgccaa 1712
OY 1561 TTGGACCTGGGACCAAGTGGCGAGGTGCTACGTGGCACTTCTGCTCCCAACCAAGC 1620
Db 1713 ttggaacctgggaaccaagtgcgagtgccacagctggaatctcgcgcacacccaagc 1772
OY 1621 GAGGCTGAGCATGAGCAGCTGACCAACGCTGAGTGAAGAACTCTAGGGCTGTGTGA 1680
Db 1773 gggggctcgagatcagcagctgacaacgctgctggaagctcctcctgagcgctgtgtga 1832
OY 1681 ACTACTGAGGCTGTCAATCATCATGGGCTAAATTCGTCAAGAAACATGCTGCGCAAGG 1740
Db 1833 actactaaggtgtcagatcatcatcagcttaactctgtcaaaagaaacatggctggaag 1892
OY 1741 GCTTCTCTTCTGGGTGTGGCTAGACAAATATCATCGACCTGTGAAAAAGATATCTTGG 1800
|||||

Db 1893 gcttcctactgggtcgtcgtagacaatalatcagacctgtgtaaaagatatacttgg 1952
OY 1801 CCTTTGAATGAAGGCTACATCATGCGCTTTCATCAGCAAGAGCGGAGCGGCATCC 1860
Db 1953 ccccttgaaagaagaggtcatcatcatatggttcatcatcagaadagagcgaggccatct 2012
OY 1861 TAAGCACAAGCCCCGCGCACCTTCTTACTGCGCTTACGCGAGAGCAACAAAGAGG 1920
Db 2013 tgaagcaactaagcccccaagcacccttcctcgtcgtctcagtgtaagcagcaagaagag 2072
OY 1921 GGGTCACTTTCATCTGGGTGGGAAAGACATCAGTGGGAGAACCCGATCCAGTCTAG 1980
Db 2073 ggtcaacttcaacttgggtggagagacatcagcgtlaaagccagatccagtlcgtgg 2132
OY 1981 AGCCATACCAACAGCAGCAGCTGAACATCATGCTCATTTGCTCAATTCATGAGCTTA 2040
Db 2133 aacctatacaagaagagcagctgaaacaacatgtaacttgcctgaaatcatatggtcata 2192
OY 2041 AGATCATGATGCGACCAACATCTGTGTCTCCACTTTGCTCACTTCACTTCACTTCC 2100
Db 2193 agatcatgtatgtaccataatctcgttgcctcactgctcatctcatctcctgacatlc 2252
OY 2101 CCAAGGAGAGGCACTTTGGAAGTCACTGAGGCCCGAGCCGAGAGACCCCGAACCG 2160
Db 2253 ccaagagagagcattcgggaagatltgctggcaagagccaagagacatcctgaagctg 2312
OY 2161 ACCGAGTAGTGTCCGCCGCTACCTGAAGACCAAGTTCATGCTGTGACACCAAGACT 2220
Db 2313 acccaagtagcgtcgtcccatatacctgaaagacaaagtatctcgtgtgacaccaaagct 2372
OY 2221 GCAGCAATACCATTTGACCTGCCGATGTCGCCCGCACCTTAAATTCATTTGATGCA 2280
Db 2373 gcagcaatacattgacctgcgtcgatgctcccggtcttagattcatltgtagcttgg 2432
OY 2281 GAAATACGGTGAAGCTGCTGAGCCCTCAGAGAGAGGGGACATTTGACGTCCCTCA 2340
Db 2433 gaaataatgtgaaagtgctgaaccccaagagagagcaaglttgaagttccctcacttg 2492
OY 2341 ACATGATCATCCTCGAGTGTGTACTCTCCCATGTGAGAGACTGAACCAAGAGCT 2400
Db 2493 acatggaatgtgacctcggagtgagctactccccaatgtgagggctgtgaagaggaagct 2552
OY 2401 GC----AGAGAGTGACTTTGAGACACCTGCCCCGTGCTCCACCCCTTAAGCGGAACC 2456
Db 2553 gcagaaagatacagctgagcgctcactgcatctgcaccccccaacagcaaaccc 2612
OY 2457 CATATGCTGAATACTCTACTTGTGTGTCGACATTTTTTTTTTAATTTCTACTTC 2516
Db 2613 cagatcatctgaaactactaactlttggtlccaga--tlttlttaactcctacttc 2669
OY 2517 TGTATCTTTGGCAATCTGGCACTTTTAAAGAGAGAAATGATGAGTGGGTGAT 2576
Db 2670 tgcatactcttgagaactctggcactctttaaataatagaagaatgagatgtggtggtat 2729
OY 2577 AAACGTTATGTAA 2591
Db 2730 ctgcttatactaa 2744
RESULT 7
AAC93150
ID AAC93150 standard; cDNA: 2787 BP.
XX AAC93150;
AC
XX
DT 15-FEB-2001 (first entry)
XX
DE Human STAT3 nucleotide sequence SEQ ID NO:1.
XX
KW Human; mouse; STAT3; phosphorothioate; antisense oligonucleotide;
KW modulation; signal transducer and activator of transcription;
KW DNA-binding; protein; signal transduction; inhibition; apoptosis;
KW inflammatory disease;; cancer; antiinflammatory; antirheumatic;

KW	cytostatic; immunostimulatory; rheumatoid arthritis; leukaemia;
KW	myeloma; melanoma; lymphoma; diagnosis; ss.
XX	
OS	Homo sapiens.
XX	
XX	WO200061602-A1.
XX	
PD	19-OCT-2000.
XX	
PF	06-APR-2000; 2000WO-US09054.
XX	
PR	08-APR-1999; 99US-0288461.
XX	
PA	(ISIS-) ISIS PHARM INC.
XX	
PI	Karras JG;
XX	
DR	WPI: 2000-619223/59.
XX	
PT	New antisense compound for inhibiting the expression of signal
PT	transducer and activator of transcription 3 (STAT3) in cells or tissues
PT	and treating diseases or condition associated with STAT3, such as
PT	rheumatoid arthritis and cancer -

Example 2; Page 74-75; 104pp; English.

The present invention describes an antisense compound (I), 8 to 30 nucleobases in length, that is targeted to a nucleic acid molecule encoding STAT3 (Signal Transducer and Activator of Transcription) and which inhibits the expression of it. (I) has antiinflammatory, antineoplastic, cytostatic and immunostimulatory activities. (I) is used for inhibiting the expression of STAT3 in cells or tissues, treating an animal having a disease or condition associated with STAT3 or a human having a disease or condition characterised by a reduction in apoptosis, and inducing apoptosis in a cell. Diseases or conditions that are treated are rheumatoid arthritis, cancer of the breast, prostate, brain, head and/or neck, leukaemia, myeloma, melanoma or lymphoma. (I) can also be used for diagnostic methods in detecting and determining the role of STAT3 in various cell functions, physiological processes and conditions and for diagnosing the conditions associated with expression of STAT3. (I) can be used alone or with other drugs as an immunostimulant. (I) is used in sandwich and colourimetric assays, involving enzyme conjugation and radiolabeling and is used in diagnostic kits. AAC93150 encodes human STAT3 and AAC93231 encodes mouse STAT3 as given in the exemplification of the present invention. AAC93151 to AAC93230 and AAC93232 to AAC93299 represent STAT3 phosphothioate antisense oligonucleotides, and AAC93300 represents a mismatch control oligonucleotide which are used in example from the present invention.

Query Match	74.7%;	Score 2144.2;	DB 21;	Length 2787;
Best Local Similarity	90.9%;	Pred. No. 0;		
Matches 2305; Conservative	0;	Mismatches 223;	Indels 7;	Gaps 2;

OY	61	GCAGCAGGAGGCTCAGTGGAAACCACTTCAGAGCTGAGACAGCTTCGGAAGCAC	120
Db	213	gaacacgagatgagcccaatggaaatcagcttacaagaagcttgaacaaggtatcctggagcaagc	272
OY	121	TGCACACAGCTGTACACAGCACACGTTTCCCATGGACCTGGCGCATTTCTCGGCACCTTGG.	180
Db	273	tccatcagctctcaagtgcagcttcccaatggagctgagcagttcttggcccccttga	332
OY	181	TTGAGAGTCAAGACTGTGGCATATGACGCCAGCAAGAAGTCACATGCGACCTTGGTTTTC	240
Db	333	ttgagagtcacaagcttgggcatactgagccagcaaaagaaatcaacatgcacacttggttttc	392
OY	241	ATTATCTCTTGGGTGGAATTTGACACCAATATTTAGCGGATTTCCGCAAGAGTCCCAATGTC	300
Db	393	ataatctccctgggagagagattgaccacagatataagccgcttcctgcgaagagctcgaaatttc	452
OY	301	TCATATGCACACCTTTCGAGAATTAAGCAGTTCCTTCGACAGACAGTATCTTGGAGAACG	360

Db	453	lctatcagcaaatctctcagaagaatcaagcagtlctcttcagagcagatcttctgagaagc	512
Db	361	CAATGGAATTTGGCCCGGATCTGTGGCCCGATTTGGCTGTGGGAAGAGTCTTGCCCTCCACG	420
QY	513	caatggagaatttgcctcggaatctgtggtcccggtgtcctgttggagaagaatcaacgctctctacaga	572
QY	421	CGCGACGCCACCGCAGCCCGACGAAGGGGGCCAGGCCAACCCCAACAGACGCCCGCTACTGA	480
Db	573	ctgagcccatctgcgcccgagcaagggggccaggaaccaccaccacacgaacgaacgctgtgtga	632
QY	481	CAGAGACGACGACAGATGTTGGAGCACCATTCTTACAGATGTCCGGAAGCGAGTGGAGGATC	540
Db	633	cggagaaagcagcaaatgtctgtgagcagacccttcgaatgtgtccggaagagatgtcagaatc	692
QY	541	TAGAACACAAAATTAAGCTGTGGAGAACCTCCAGAGCAGCTTTGATTTCATACATCAAAAA	600
Db	693	tagaaccaaaaaatgaagaatgtgtatagagaatcccaagatgtactttagatcttcaactataaa	752
QY	601	CCCTCAAGAGCCGAAGACATCGAGATCTGATATGGAACCAACCAAGTCTGTACACGAC	660
Db	753	ccctcaagaatgcaagagacacatgcaagatcttgaatgtgaacaacacagatgtgacgaagc	812
QY	661	AGAATATCGACGACCTGTGAGACATGTCACAGCCCTGGACACAGATGGCGGAAGCATTTG	720
Db	813	agaagaatgcagcaacgttgaaacaagatgtcctacgtcgtctgtgacaagaatgtcgagaagcatcg	872
QY	721	TGAGTGACCTGGCGGGGCTCTTGTACAGCAATGGAGTACTGTGAGAAGACACTGACTGATG	780
Db	873	tgaatgtacgtctgtgaggggtctttgttcagcgatgtgagtacgtgcagaaaaactctcaagcgcg	932
QY	781	AAGAGCTGCTGTACTCGACAGAGCGCGCCAGAGATCGCGTGTGATCGGAGGCGCTCCCAACA	840
Db	933	aggaagctgtgcgtacgtgaagggcgcgaaacaagaatgtacctgtcatgtgagggcccgcaaca	992
QY	841	TCTTGGC. _ACCGTCTGSAAAACGTGATACCTTATTACAGAACTCTCAACTTCAGACC	900
Db	993	ctgtgcctgatctcgtgtagaaaacgtgaaacgtcatatgacgaatcccaactcagaacc	1052
QY	901	GCCACACAAATTAAGAAACTGTGAGAGAGTGTGACGAGAAAGTGTCTTACAAAGGCGACCTA	960
Db	1053	gtcaacaataatlaagaagaacttgaagagtgtgaccaaaaagtlctctacaagaaggagcccca	1112
QY	961	TGCTGCAGACCGCGCCATGCTGTGAGAGAGAGATCTGTGAGAGCTGTTCACAAACTTAATGA	1020
Db	1113	ctgtacagagaccgcgatgtctgtgaagagagatctgtgagctgtgtctcagaataactlaatga	1172
QY	1021	AGAGTGCCTTGTGTGTGTGAGGCGGAGCGCTGCATGCGCATGACGCCGAGCGGCTTAA	1080
Db	1173	aaagtgcctctgtgtgtgtgagcgagacccgtcgaatgcacatctctgtacggtccggtccctcg	1232
QY	1081	TGATCAAGACTGTGTCTCCAGTTTACCGAGAAAGTCAAGTGTGCTGATCAAAATTTCTTGAGT	1140
Db	1233	tcatcaagaacggtgtccagttacactactactaaagtlcaagtlgtctgtgtcaagtlctcctgagt	1292
QY	1141	TGAATTTATCAGCTTAAATTTAAAGTGTGCATGTATAAAGACTGTGGGAGATGTTGCTGCC	1200
Db	1293	tgaattatcagctttaaataatlaaagtgtgtcatgtlgaacaaagactcttggggacgtttcgaatc	1352
QY	1201	TTCAGAGGCTCTCGGAATTTTAACATTTCTGGGACGAAACAAAAGATGAAACATGTGAGG	1260
Db	1353	tcaagagatctccggaataatlaacatctctgtggcaacaaacacaaaagtgtgtgaaca-vgaaag	1412
QY	1261	AGCTTAAACAACGGCAGCTGTCTCTCAGAGTGTAAAGCACTGACCTTTAGGGAGCAGAGAT	1320
Db	1413	aatccaacaacagcgactctctctgcagaatltcaaaacacttgaccctgtgagggcagagat	1472
QY	1321	GTGGGAATGAGGCGCGTGCACAAATTTGTATGTCTCTTTGATCTGTGACTGTAGAGACTTGAC	1380
Db	1473	gttgggaatgtggggcgagcgacaattgtgatgtcttccctgtattgtgactgtgagggctgtcac	1532
QY	1381	TGATCACTTGCAGACTGTAGGTGTACCAACCAAGGCTCTCAAGATTTGACTTAAGACCCACT	1440

Db 1533 tgcataccttggagaccgaggtgltatcaaccaaggtctcaaatgtgacctgaagaccact 1592
Qy 1441 CCTTGGCAGTTGTGTGATCTCCCAACATCTGTGAGATGCCAAATGCTTGGCATCAATCC 1500
Db 1593 ccttgcagtggtgtgtagatctccaacatctgtcaagatgccaaatgcttggcgtccatcc 1652
Qy 1501 TGTGTATTAACATGCTGACCAATTAACCCCAAGAACTGAACTTCTTACTTAAGCCGCAA 1560
Db 1653 tgytgaacaacatgctgacaacaatcccaagaatgtgaactcttcaactaagacgcgcaaa 1712
Qy 1561 TTGGAACCTTGGGACCAAGTGGCGAGTGTGTCAGCTGGTCTGCTCCACCCCAAGC 1620
Db 1713 ttggaacctgggaccagtggtccgaggtgtccagcttgcgtccatcccaaccaagc 1772
Qy 1621 GAGGCGTGAACATCAGACAGCATGACAGCGTGGCTGAGAGCTTCTAGGGCTTGGTGA 1680
Db 1773 gggggcctgagcatcgagcagctgacaacagctgtgctgagaagctccctagggcctgtgtga 1832
Qy 1681 ACTACTGAGGGTGTGAGATCAGATGGGCTAAATTTGCAAGAAAGAAACATGGCTGGCAAGC 1740
Db 1833 actactcaaggtgtcagatcacatggtctaactctgcaagaacacatgctgtgcaag 1892
Qy 1741 GCTTTCCTTCTGGGTCGTGGCTGACATATATCATCGACCTTGTGAAAAGATATCTTGG 1800
Db 1893 gcttctctctacgtggtctgtgctagacaatatcatcgaccttggaaaaagatactcttgg 1952
Qy 1801 CCCTTTGAATGAAGGGTACATCATGGGTTTCATCAGCAAGAGAGCGGAGCGGCCCATCC 1860
Db 1953 cccttggaaatgaaggttatcatcatctggttctcatcagcaagagcggtgagccatctc 2012
Qy 1861 TAAGCAACAAAGCCCCCGGCGACCTTCTACTGGGCTTACGCGAGAGCAGCAAGAAAGAGAG 1920
Db 2013 tggagcaaaagcccccaagagaccttctgtgctgtcagctcaatgaaagcagcaaaagagag 2072
Qy 1921 GGGTCACTTTTCACTTGGTGGTGAAGAGACATCAGTGGAGACCCAGCATCTGTAG 1980
Db 2073 ggtctacttctacttgggtgtgagaagacatcagcgttaagaccagatccagtcgttgg 2132
Qy 1981 AGCCATACCAACAGACAGCAGCTGAACAACATGTCAATTTGCTGAATCATGATGGCTATA 2040
Db 2133 aaccatacaacaagcagcagctgacaacatgtcatctgtgaatcatcatctggtgctata 2192
Qy 2041 AGATCATGATGGAGCAACATCTGTGTCTCCACTTGTCTACTCTTACCCGACATTC 2100
Db 2193 agatcctgtgtgtaccataatccctgttcttccactgtctatctatctatctgacatc 2252
Qy 2101 CCAAGAGAGAGGAGGATTTGAAGTACTGTAGGCCGAGAGCCAGAGCCGCAAGCCG 2160
Db 2253 ccaagagagagagcattcgtggaagatattgtgcagagagagccagagcattcgtgaagctg 2312
Qy 2161 ACCGAGGATGCTGCCCGCTACCTGAAGACCAAGTTTCATCTGTGTGACACCAAGCAGCT 2220
Db 2313 acccaggttagcgtcccccatactctgagaagcagttatctgtgtgacaccaaagcct 2372
Qy 2221 GCAGCAATCCATTAAGCTGCGATGTCGCCGATGCCCGCACTTAAATTAATTCATGATGAGTTTG 2280
Db 2373 ggcagcatatcatctgacccgtgcgaatgccccccggtcttagatctatgtatgcaagttg 2432
Qy 2281 GAAATTAAGGTAAGGTGCTGAGCCCTGACAGCAGAGAGGAGGAGTTGATGCTCAGCTTTG 2340
Db 2433 gaataatgtgtgaaggtgtcgaacccctcagcagaggggaggttggctccctcaacttg 2492
Qy 2341 ACATGATATGACCTTGGAGAGTGTGTACTCTCCCATGTAGAGAGCTGAACCAAGAGCT 2400
Db 2493 acctgtgagttgacctgtgaggtgtcgttaacctcccccatgtlgaagagcgtgaaaggaagct 2552
Qy 2401 GC----AGAGACGTGACTTGAGACACCTGCCGCTGTGCTCAACCCCTTAAGAGCGGCAAGCC 2456
Db 2553 ggcgaagaagatacagactgagagcgccctacgtgcatcttgcceccctcaacagcgaacccc 2612
Qy 2457 CATATGCTGTAAGATCTCAACTTTGTGTGATTCAGATTTTATTTTATTTCTACTTTC 2516
Db 2613 cagatcatctgaaactaactaacttgtgtgtcaga---tttttttaactctctacttc 2669

Qy 2517 TGTATCTTTGGGCAATCTGGGACACTTTTAAAGAGAAATGATGATGTGGTGCAT 2576
Db 2670 tgcataccttggagcaatctgggacattttaaataagaaatgaaatgtgagtgtggtgat 2729
Qy 2577 AAATGTTATGTATA 2591
Db 2730 ctgcttltatcta 2744

RESULT 8
AAA89229 standard, cDNA: 2787 BP.
ID AAA89229:
AC AAA89229:
XX
DT 28-MAR-2001 (first entry)
XX
DE Human signal transducer and activator of transcription STAT-3 DNA.
XX
KW STAT-3; signal transducer and activator of transcription 3;
KM human; crystal; drug screening; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 221..2533
FT /*tag= a
XX
PN U56160092-A.
PD 12-DEC-2001
XX
PF 29-MAY-1998: 98US-0087465.
XX
PR 29-MAY-1998: 98US-0087465.
XX
PA (UVRQ) UNIV ROCKEFELLER.
PI Chen X, Darnell JE, Kuriyan J, Vinkemeier U, Zhao Y, Jeruzalmi D;
XX
DR WPI: 2001-101568/11.
DR P-PSDB: AAB19964.
XX
PT Novel crystal useful in drug screening assays, comprises portion of
XX signal transducer, activator of transcription and duplex DNA
XX
PS Disclosure: Column 63-68; 206bp; English.
XX
CC The present sequence is that of DNA encoding human signal
CC transducer and activator of transcription 3 (STAT-3, see AAB19964).
CC The invention provides a crystal of a core portion of a STAT
CC protein in dimer form with an 18-mer duplex DNA (see AAA89233) that
CC contains a binding site for the STAT dimer. The core portion
CC comprises an N-terminal alpha-helical domain, a DNA binding domain
CC which contains an immunoglobulin-like fold, a C-terminal SH2 domain
CC and a domain that links the DNA binding and SH2 domains. The
CC crystal is of sufficient quality to perform X-ray crystallography
CC studies. Methods of preparing the crystals are included in the
CC invention. Knowledge of the STAT protein's 3-dimensional structure
CC will aid in structure-based drug design. The crystal can be used
CC in drug screening assays to identify agonist and antagonist
CC compounds. Antagonists can be used to treat inflammation, allergy,
CC asthma and leukaemia, and agonists to treat anaemia, neutropenia,
CC thrombocytopenia, cancer, obesity, viral diseases, growth
CC retardation, and other conditions characterized by insufficient
CC STAT activity.
XX
SQ Sequence 2787 BP; 729 A; 719 C; 753 G; 586 T; 0 other;

Query Match 74.7%; Score 2144.2; DB 22; Length 2787;
Best Local Similarity 90.98; Pred. No. 0;

QY 2221 GCACCAATACCATTTGACCTGCCGATGTCGCCCGCACTTATGATTCATTCAGTTTG 2280
|||||
Db 2273 gcagcaatacattgacctgcgcgtgccccgcgctttagatcaattgagcagtttg 2432
QY 2281 GAATAACGCTGAAGCTGCTGACCCCTCAGCAGAGGCGAGTTGATCCGCTCAGCTTGG 2340
|||||
Db 2433 gaataatggaaggtgctgaaccctcaagaaaggaagcagttgagtcctcctcagcttg 2492
QY 2441 ACATGATCTGACCTCGAGCTGCTGCTACCTCCCATGTGAGAGCTGAACAGAGAGCT 2400
|||||
Db 2493 acatgaggttgaactctgagtgagctacccctccatgtagagagcaggaagcgt 2552
QY 2401 GC---AGAGCTGAGCTTGAGACACTGCCCGCTGCTCCACCCCTAAGCAGCGCAACC 2456
|||
Db 2553 gcagaaagatcagactgagcgcctacactgcatctgtccaccctcacagccaaccc 2612
QY 2457 CATATCGCTGACCACTCTACTGTTGTTGCCAGATTTTATTTTAAATTCCTACTTC 2516
|||||
Db 2613 cagatcactgaaactactaactctgtggtccaga---ttttttaaactcctcacttc 2669
QY 2517 TGTATCTTTGGCAATCTGGCGACTTTTAAAGAGAGAAATGAGTGTGGGTGAT 2576
|||||
Db 2670 tgcatactttagaactctgagcactttaaataagaaatgagatgagtggtgagat 2729
QY 2577 AACTGTATGTATA 2591
|||
Db 2730 ctgccttactctaa 2744

RESULT 9
AAF18318
ID AAF18318 standard; DNA; 3156 BP.
XX
AC AAF18318:
XX
DT 14-MAR-2001 (first entry)
XX
DE Lung cancer associated polynucleotide sequence SEQ ID 337.
XX
KW Human: lung cancer associated protein; neuroprotective; cytostatic;
KW cardiocactive; immunomodulatory; muscular active; vulnerary;
KW gastrointestinal; nephrotoxic; antiinfective; gynecological;
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW proliferative disorder; wound healing; infectious disease; ds.
XX
OS Homo sapiens.
XX
PN W020005180-A2.
XX
PD 21-SEP-2000.
XX
PE 08-MAR-2000; 2000WO-US05918.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
PI Ruben SM:
XX
DR WPI: 2000-587514/55.
DR P-PSDB: AAB58442.
XX
PT Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer -
XX
PS Claim 1; Page 794-795; 1425pp; English.
XX
CC polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective; cytostatic; cardiocactive;

CC immunomodulatory; muscular active general; vulnerary; gastrointestinal
CC general; nephrotoxic; antiinfective; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the
CC protein or polynucleotide sequences. The lung cancer associated
CC polynucleotide sequences may be used for detection of lung cancer,
CC chromosome identification, as chromosome markers, and for numerous other
CC diagnostic or research purposes. The proteins may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders. The proteins may also be used in the treatment of wounds and
CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
CC peptide AAB58549 are used in the course of the invention for the
CC identification and characterisation of the polynucleotide and protein
CC sequences.
XX
SQ Sequence 3156 BP; 832 A; 792 C; 826 G; 705 T; 1 other:

Query Match 74.3%; Score 2132.6; DB 21; Length 3156;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 2464; Conservative 0; Mismatches 334; Indels 51; Gaps 8;

QY 68 GATGGCTCAGTGAACCAAGCTGACAGAGCTGACACAGCTACCTGAAGCAGCTGACCA 127
|||||
Db 194 gatgcttcatgataatcagctacagcagcttgacacacagctgagcagctccatca 253
QY 128 GCTGTACAGCCACAGCTTCCCATGAGAGCTGGCGACTTCTCGACCTGGATTGAGAG 187
|||||
Db 254 gctctacagtagcagctcccaatgagcgtcggcagcttctgcccccttgagtgag 313
QY 188 TCAAGACTGGGCAATGAGCAGCAGCAAGAGTCAATGCCAGTGTGTTCTAATATCT 247
|||||
Db 314 tcaaatgagcatatgagcagcagcaagaatcacatgcaccttggttctaatct 373
QY 248 CTTCGCAATATGACCAATATAGCCATTTCTGACAGAGTGCATGCTCTATCA 307
|||||
Db 374 cctgcttcatgataatgacccagctagcttctcctcctcctcctcctcctcctcctc 433
QY 308 GCACACCTTGGAGAAATCAAGCAAGTTCCTGACAGCAGTATCTTGAAGCCCAATGA 367
|||||
Db 434 gcaacatcacaagaatcaagaagcttctcagagcagctatcttgagaagccaatgga 493
QY 368 AATGGCCGGATCTGTGCGCCATGCTGTGGAAAGTCTGCTCTCCAGCGGACG 427
|||||
Db 494 gatgcccagattgtgcccgtgcttgggaagaatcaacgctctacagactgcagc 553
QY 428 CACGGCAGCCCAAGAGGGGCGCCAGCCCAACCCAGCCCGGATGACAGAGAA 487
|||||
Db 554 cactgcccagcaagaagggcgccagggcccaaccaccaagcagcggcgtgagcggaa 613
QY 488 GCAGCAGATGTTGGAGCAGCATCTTACAGATGTCGGAAGCGAGTGCAGATCTAGAACA 547
|||||
Db 614 gcagcagatgcttgagcagcagcactcagatgctcggagagagtgcaagctcagaca 673
QY 548 GAATAATGAAGTGTGGAGAACCTTCAGAGAGCTTTGATTTCAACTCAAAACCTCAA 607
|||||
Db 674 gaaatgaaagtgtagaagaaatctccagagactttagttaaactataaaacccctcaa 733
QY 608 GAGCCAGAGAGCATGAGATCTGAATGAAGAACACACAGCTCTGACACAGAGAGAT 667
|||||
Db 734 gagtcagaagagacatgacatgcaatggaatacaaccacagcagctgacacagcagaagat 793
QY 668 GCAGCAGCTGGAACAGATGCTACAGCCCTGACCCAGATGCGGAGAAAGCATTTGAGTGA 727
|||||
Db 794 gcagcagcttgaaacagatgctcactgctgctgacacagatgcggaagacatcagtgta 853
QY 728 GCTGGCGGAGCTCTTGTACAGATGATGATGAGAGAGACTGATGATGAGAGCT 787
|||||
Db 854 gctgcccggagcttgtcagcagatgagtagcagtgcaagaacactcagcagcagagcct 913
QY 788 GCGTCACTGGAAGAGCGCGCAGAGATGCTGATCGAGAGCCCTCCACATCTGCTCT 847
|||||
Db 914 gctgactgaggaagggcgcaacagatgctcctgacttgagagccgcccacacatctgct 973

OY	1201	TCGAGGCTCTCGGAAATTTAACTTCTGCGCACGACAAAGTGATGATACTGAGG	1260
Db	1353	TCAGAGGATCCCGGAAATTTAACTTCTGCGCACGACAAAGTGATGATACTGAGG	1412
OY	1261	AGCTCTAACCAAGGACACCTGTCTGCGAGATTCTAAGCAGACCTCTTAGGAGCAGAT	1320
Db	1413	AATCCACAAGGACGAGCTCTGCGAGATTCTAAGCAGACCTCTTAGGAGCAGAT	1472
OY	1321	GTGCGAATGAGAGCCGTGCCAATTGTGATGCTCTTGTGATCTGACTGAGAGCTGCACC	1380
Db	1473	GTGCGAATGAGAGCCGTGCCAATTGTGATGCTCTTGTGATCTGACTGAGAGCTGCACC	1532
OY	1381	TCATTCACCTTCGAGACTGTGAGGTTACACCAAGGCTTCAAGATTGAGCTGAGAGCCACT	1440
Db	1533	TCATTCACCTTCGAGACTGTGAGGTTACACCAAGGCTTCAAGATTGAGCTGAGAGCCACT	1592
OY	1441	CCTTGCAGATTGTGTGATCTCCACACATCTGTGAGATGCCAAATGCTTGGGCATCAATCC	1500
Db	1593	CCTTGCAGATTGTGTGATCTCCACACATCTGTGAGATGCCAAATGCTTGGGCATCAATCC	1652
OY	1501	TCGTGATTAACATGTCGACCAATAAACCCCAAGAACGTGAACTTCTACATAAGCCGCCAA	1560
Db	1653	TCGTGATTAACATGTCGACCAATAAACCCCAAGAACGTGAACTTCTACATAAGCCGCCAA	1712
OY	1561	TTGGAACTCTGGGACCAAGTGGCCGAGGTGCTGACGTGGCAGTTCTGCTCCACGACCAAGC	1620
Db	1713	TTGGAACTCTGGGACCAAGTGGCCGAGGTGCTGACGTGGCAGTTCTGCTCCACGACCAAGC	1772
OY	1621	GAGGCGTGAACATCGAGACACTTACAACGCTGGCTGAGAAAGTCTTAGGGCTGTGTGA	1680
Db	1773	GAGGCGTGAACATCGAGACACTTACAACGCTGGCTGAGAAAGTCTTAGGGCTGTGTGA	1832
OY	1681	ACTACTCAGGGTGTGACATCACATGGCGTAAATTCTGCAAGAAACATGAGTGTGCAAG	1740
Db	1833	ACTACTCAGGGTGTGACATCACATGGCGTAAATTCTGCAAGAAACATGAGTGTGCAAG	1892
OY	1741	GCTTCTCCTTCTGGGCTGTGACTAGACAATTCATCGACCTTGTGAANAAGTATATCTTGG	1800
Db	1893	GCTTCTCCTTCTGGGCTGTGACTAGACAATTCATCGACCTTGTGAANAAGTATATCTTGG	1952
OY	1801	CCCTTTGGAATGAAGGTATCATATGCGTTTCATGCAAGGAGGCGGAGCGGCGCATCC	1860
Db	1953	CCCTTTGGAATGAAGGTATCATATGCGTTTCATGCAAGGAGGCGGAGCGGCGCATCC	2012
OY	1861	TAAACACAAAGCCCGCGGACCTCTCT	1888
Db	2013	TAAACACAAAGCCCGCGGACCTCTCTCT	2072
OY	1889	-----ACTGGGCTTACCGAGACGACGAAGAAGAGAG	1920
Db	2073	-----ACTGGGCTTACCGAGACGACGAAGAAGAGAG	2132
OY	1921	GGGTCACTTTCACCTTGGGTGGAAAAGACATCAGTGGCAAGACCCAGATCCAGTCTTAG	1980
Db	2133	GGGTCACTTTCACCTTGGGTGGAAAAGACATCAGTGGCAAGACCCAGATCCAGTCTTAG	2192
OY	1981	AGCCATACACAAAGCAGACGACTAACACACATGTTCATTTGCTGAAATCATCATGGGCTATA	2040
Db	2193	AGCCATACACAAAGCAGACGACTAACACACATGTTCATTTGCTGAAATCATCATGGGCTATA	2252
OY	2041	AATATCATGATGCGACCAACATCTGTGTCTCAGTGTCTACCTTACCCGACATC	2100
Db	2253	AATATCATGATGCGACCAACATCTGTGTCTCAGTGTCTACCTTACCCGACATC	2312
OY	2101	CCAAAGGAGAGGCAATTTGGAATGACTGTAGGCCGAGACCCAGAGACACCCGACACCG	2160
Db	2313	CCAAAGGAGAGGCAATTTGGAATGACTGTAGGCCGAGACCCAGAGACACCCGACACCG	2372
OY	2161	ACCCAGGTAGTGTGCGCCCTACCTGAAGACCAAGTTCACTGTGTGACACCAACACACT	2220
Db	2373	ACCCAGGTAGTGTGCGCCCTACCTGAAGACCAAGTTCACTGTGTGACACCAACACACT	2432

QY	2221	GCACGACATACCATTTGACCTCCGACATGTCCTCCCGGACCTTACATTTGATGACGTTTG	2280
Db	2433	gcagaaataacatctgacctccgcagatgctcccccgcgcttgatcatcattgaagcaagttt	2492
QY	2281	GAATTAACGCGAAGCGTGCAGACCCCTACACAGAGAGGAGATTGTAGTCGCTACGTTTG	2340
Db	2493	gaataataagtggaagagtgctcgaaacccctcagacagagagggcagtttgagttccctaccctt	2552
QY	2341	ACATGGATCTTGACCTCGGAGCTGTCTACGTCCTCCCATGTGAGAGACTGAAACACAGACT	2400
Db	2553	acatgagatgacctcgcgagtgctgcctaccctcccatctgagagagcttgagcaagaaact	2612
QY	2401	GC----ACAGACGCGACTTGAGACACTGGCCCGCTGCTCCACCCCTAACACACGCAACCC	2456
Db	2613	gcagaaagaatactgacgcgaagcgccctaccctgcatctgcacccctccacacagcaaaccc	2672
QY	2457	CATATCGCTGACAACTCGTAACCTTGTGGTGCCTGCAGATTTTATTTTAAATTCCTACTTC	2516
Db	2673	cagatcatcgcgaacatcaataccttctggtgtccaga---tttttttaatccctaccctc	2729
QY	2517	TGCTATCTTTTGGGCAATCTGGGCACTTTTAAAGAGAGAAATGACGTAGTGTGGTTCAT	2576
Db	2730	tgctatcctcttgagcaatcctgcgacacttttaaaatagagaatgagtgatgagttgagat	2789
QY	2577	AAACTGTTATGTTAA	2591
Db	2790	ctgctttatactaa	2804

RESULT 11

ID AAX29281 standard; cDNA; 2344 BP.

AC AAX29281;

DT 11-JUN-1999 (first entry)

DE Human STAT3 allelic variant encoding cDNA.

KW Signal Trans: Inter and Activator of Transcription 3; STAT3; allele; IL-6

KW pharmaceutical; autoimmune disease; inflammatory; human; ss

OS Homo sapiens.

PN EP905234-A2.

PD 31-MAR-1999

18-FEB-1998; 98EP-0102774.

16-SEP-1997; 97EP-0116061.

PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.

PI Della Pietra L, Ser

WPI: 1999-192664

F-730B, 65103/00.
DN
XX

PT New human allelic

p1 disease
XX XX

XX CLALL 3

CC The present sequence represents a predominant allele
CC variant of human Sigval Transducer and Activator of Transcription 3

CC (SWHS) protein, an intracellular transcription factor which mediates
CC IL-6 signals. This sequence differs from the original published human

CC recombinant STAT3 DNA molecule can be used for the recombinant expression

OY	305	TCGACGAACCCCTCCAAAGATTAACAGACAGTTTCTGCAGACAGAGTTCCTTGACAAACCCAT	364
Db	241	gcagcacaacatcagcgaagaagcagcgtatctcccaagataactctccaaagaagctccgt	300
OY	365	GGAAATGTGCCGATCTGTGGCCCCGATGCCCTGTGGAGAGAGTTCCGCTCTCCACAGCGC	424
Db	301	acagatgtccatgcatcatctacaaactgtctcgaagaaagaagaagatcttggaaatgtc	360
OY	425	AGCCACGGCAGCCCGCAAGGGGGCGAGCCCAACCAACCCACAGCGCCGCTGTGTGCACA	484
Db	361	ccaaagatttaatca-----ggcccaggaaggaaatctcgaacacactgtatgttaga	414
OY	485	GAAAGACAGATGTTGGACACACATCTTACAGATGTCCGAAAGCCAGTGTCCAGATCTACA	544
Db	415	taaacagaagcgtctgacagtaaagctcagaataatgtgaagagatcaagtcattgtcataga	474
OY	545	ACAGAAAATGAAGGTGTGTGGAGAACCTCCAGAGACGACTTTGATTTCATACACAAAACCTT	604
Db	475	gcaggaataatcaagaccctgaagaatctaacagaatgataatgtacttaaatgtcnaaacctc	534
OY	605	CAAGAGCCCAAGGAGACATCAGAGATCTGAAATGGAAMCAACCAAGCTCTGTACCCAGACAGA	664
Db	535	tcagaaacaagaag-----gtgaagccaatgtgtgtgcgaagaagcagcaaaaacaaga	588
OY	665	GATGCGAGCTGTGAACACATATGTCACACCCTCGAGCCAGATCCGAGAAAGCATTTGTGAG	724
Db	589	acagctgcgtctccacaagatgttttaatgtcttgacataagaagaagataatctca	648
OY	725	TGAGCTGGCGGGCTTTGTTCAGCATATGAGATACGTGCAGAGACACACTGACTGATCAAGA	784
Db	649	caaaatcagaagatgtctgaatctcatctcagctcactcagaacaacctgtatctaaagca	708
OY	785	GTGCGCTGACTGTGAAGAGCGCGCCAGACATCGCGTCCATCGAGCGCCCTCCAAACTCTT	844
Db	709	gtccgtgtgagtgtagaagcgaagcagcagcagcgtcatcggtgagcgcccaagcgtg	768
OY	845	CCTGGAACCGTCTGAAAACGTGATACTTCATTATAGCAAGATCTCAACTTCAGACCCGCCA	904
Db	769	ctctggaatcagcttgcacaaagctgtgttccaccattgttgcagaagaccctgcagaagatccgta	828
OY	905	ACAAATTAAGAAACTGAGAGAGCTGCAGACGAAGAAGTCTCTCAAGCGCACCCCTATCGT	964
Db	829	gcagcttaaaagcttgagaggtgttgaaacagaataatcactatgagccgcgaacctattac	888
OY	965	GCAGCACCGGCCCATGCTGTGAGAGAGAGATCTGTGAGACTTTTCAGAAACTTAATGAAGAG	1024
Db	889	aaaaaacaagcaggtgtgtcagatcgaacctctctctctccagaagatcatcgaag	948
OY	1025	TGCGTTTCGTGTGAGCGGAGCGCCCTGCATGCCCATGCACCGAGCCGACTTATGATCAT	1084
Db	949	ctctctcgtgtgtagaagcagaagcgtgtgcatgcacacacccgcagagagccctgtgctc	1008
OY	1085	CAAGACTGTGTGCCAGTTTACCAAGAAAGTCAGGTGCTGTGCTCAAAATTTCTCAGTTGAA	1144
Db	1009	gaagactgtgtgtacagttcacctgtcgaagtcgaagctgtgtgtgaataatcgaagatctgaa	1068
OY	1145	TTATACACTTAATAATTAACTGTGCATTTGATTAAGAAGACTCTGGGAGATTCTCTGCCCTCAG	1204
Db	1069	tcctattaagaaagtgaaatgtcactcttgccaagatgtgaacgagagaaaaacacagttaa	1128
OY	1205	AGGCTCTCGGAAATTTTAACATTCTGGGCACGAAACAAAAGTGTATGAACATGGAAGAGTCT	1264
Db	1129	aggaatttcggaadltcacactcttggtaagcacaacaaagtatgatacatgagaagatc	1188
OY	1265	TAAACAAGCAGAGCTGTCTCAGAGATTCAAGCAACCTGACCCCTTAGAGAGAGAGATGTGG	1324
Db	1189	caccacaagcagatctgtgcagctgagctccgaacacctgcaactgaaagaaacaagaa-----	1243
OY	1325	GAATGAGGCGCTGCCAATTGTGATGCTCTCTTGATGTGACTGTAGAGAGCTGCACCTGAT	1384
Db	1244	-aaagctgtggaacagaactaattagaggccctctaatgttcatccgaagaactctcctc	1302

QY	1385	CACCTTGAGACTGAGGTGTACACCAAGGCGCTCAAGATATGACCTAGAGACCCACTGCTT	1444
Db	1303	tagctcttgaaccacagctctgtccacgaagccttctgtatcttgacccttgagaccactctct	1362
QY	1445	GCCAGTTGTGGTGTATCCCAACATCTGTGAGATGCCAAATGCTTGGCATCAATCTGTG	1504
Db	1363	tccgtctgtgtgtatctcccaagctccagcgaagctccccaagtgtcgtgtgtgtatctacgtg	1422
QY	1505	GTATACATGCTGACCAATTAACCCCAAGAACGTGAACCTTTTCACTAAGCCACCAATTGC	1564
Db	1423	gtacaacatgtctgtgcacagaccaggaatctctctctctctcttgaaaccccccgtctgc	1482
QY	1565	AACCTGGAGACCAAGTGGCCCGACAGTGTCTACACTGGCAGTTCTGTGTACCACCAAGCCAG	1624
Db	1483	gtgtgtgtgtccacagctctcaagatgtgtatgtgttgcagttctatcaagtcacccaagaag	1542
QY	1625	GCTGAGCTTCAGAGAGCTGCAACGGCTGGCTGAGAGCTCTTAGGGCGCTGGTGTGACTA	1684
Db	1543	tctgaagcgaacacagcctgtgacatgtctgtgagagaagcctgcctgtgccttaatgc----	1599
QY	1685	CTCAGGCTGTGAGATTCACATGGGCTTAATTCTGCAAAAGAACATGGCTGGCAAGGCGTT	1744
Db	1600	cccgtatgtcttctatcccatggaacaggtctttagaagaaatatatgatlaaataatlt	1659
QY	1745	CTCCTCTTGSGGTCTGCGCTAGACATATCATCTGCAGACCTTTGGAAAAGTATATCTTGCCCT	1804
Db	1660	ctctctctctgtgccttctgtatctgacacatctccatagagctctatlaagaacagacctgtgtgc	1719
QY	1805	TTGGAATCAAGGGGTACATCATGAGGTTTTCATCACCAAGAGACGGAGCGGCGCATCTTAAG	1864
Db	1720	ctggaatatgtatgtgtcatctatgtgtcttcatcagacaaggtcgagaaagcgtctgtctcaa	1779
QY	1865	CACAAAGCCGCCGGGACCTTCTACTGCGCTTTCAGCGAGAGCAGCAAAAGAGAGGCGT	1924
Db	1780	ggaaccagcagccagaagcgttctctgtcttgatctcagtgtagagcgtcccggaagggccat	1839
QY	1925	CACCTTTCACCTGGGTGCGAAAAGAC---ATCACTGGCAGAGACCCAGATCACTGTGTGA	1981
Db	1840	caactctcatatgtgtgtgaagcgtgtcccaagaagaggtgaaacctgtacatctcatatgcgtgga	1899
QY	1982	GCCATACACCAAGCGACGCTGGAACAACATGTATATTGGTGAATCATCATGCGCTATTA	2041
Db	1900	gacctacaagaaaaagaactctttagctgttactcttccagatatatcttcgaactaca	1959
QY	2042	GATCATGTAGTGGAGCAACATCTGGTGTCTCCACTTGTCTACCTCTACCCCGACATTCC	2101
Db	1960	agtcatagtctgcggagacataccagaagaatccctgaagtatctgacccaatatgtga	2019
QY	2102	CAGAAGAGAGCATTTTGGAAAGTACTGT	2129
Db	2020	caaaagacacagccttctgtggaatatat	2047

RESULT	14
AAC98863	
ID	AAC98863 standard; cDNA; 2974 BP.

DT 09-MAR-2001 (first entry)

DE	Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:91.
AA	

Human; pancreas; pancreatic cancer; pancreatic cancer antigen; detection; diagnosis; identification; cytosolic; neuroprotective; neoplastic; immunomodulatory; relaxant; contraceptive; gynecology; antiinflammatory; cardiac; gene therapy; chromosome mapping; linkage analysis; tissue identification; tissue typing; forensic; immune system; muscular; reproductive; gastrointestinal; neural; cardiovascular; renal; proliferative; ss.

OS Homo sapiens

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2002, 14:00:58 ; Search time 63.17 Seconds
(without alignments)
10285,980 Million cell updates/sec

Title: US-08-212-185-11

Perfect score: 2869
Sequence: 1 GCCGCCGACGCCAGCCGCCG.....AATTAAAAAAAAAAAAA 2869

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/1na/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/1na/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/1na/PTOS.COMB.seq:*
6: /cgn2_6/ptodata/2/1na/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2869	100.0	2869	1	US-08-369-796-11 Sequence 11, Appl
2	2869	100.0	2869	2	US-08-852-091-11 Sequence 11, Appl
3	2869	100.0	2869	2	US-08-820-754-11 Sequence 11, Appl
4	2869	100.0	2869	3	US-08-956-652-11 Sequence 11, Appl
5	2869	100.0	2869	3	US-08-956-869-11 Sequence 11, Appl
6	2869	100.0	2869	3	US-08-948-547-11 Sequence 11, Appl
7	2869	100.0	2869	5	PCT-US95-17025-11 Sequence 11, Appl
8	2864.2	99.8	2869	3	US-09-288-461-82 Sequence 82, Appl
9	2861	99.7	2869	4	US-09-364-970-9 Sequence 9, Appl
10	2415.6	84.2	2652	1	US-08-416-581B-7 Sequence 7, Appl
11	2415.6	84.2	2652	1	US-08-416-581B-8 Sequence 7, Appl
12	2294	80.0	2310	1	US-08-416-581B-6 Sequence 6, Appl
13	2144.2	74.7	2787	1	US-08-416-581B-3 Sequence 3, Appl
14	2144.2	74.7	2787	1	US-08-416-581B-4 Sequence 4, Appl
15	2144.2	74.7	2787	3	US-09-288-461-1 Sequence 1, Appl
16	2144.2	74.7	2787	4	US-09-087-465-5 Sequence 5, Appl
17	2026.8	70.6	2310	1	US-08-416-581B-2 Sequence 2, Appl
18	682	23.8	2277	2	US-08-369-796-7 Sequence 7, Appl
19	682	23.8	2277	2	US-08-852-091-7 Sequence 7, Appl
20	682	23.8	2277	3	US-08-820-754-7 Sequence 7, Appl
21	682	23.8	2277	3	US-08-956-652-7 Sequence 7, Appl
22	682	23.8	2277	3	US-08-956-869-7 Sequence 7, Appl
23	682	23.8	2277	3	US-08-948-547-7 Sequence 7, Appl
24	682	23.8	2277	5	PCT-US95-17025-7 Sequence 7, Appl
25	672.6	23.4	2607	2	US-08-369-796-5 Sequence 5, Appl
26	672.6	23.4	2607	2	US-08-852-091-5 Sequence 5, Appl
27	672.6	23.4	2607	2	US-08-820-754-5 Sequence 5, Appl

ALIGNMENTS

28	672.6	23.4	2607	3	US-08-956-652-5	Sequence 5, Appl
29	672.6	23.4	2607	3	US-08-956-869-5	Sequence 5, Appl
30	672.6	23.4	2607	3	US-08-948-547-5	Sequence 5, Appl
31	672.6	23.4	2607	5	PCT-US95-17025-5	Sequence 5, Appl
32	672.6	23.4	3943	1	US-08-369-796-3	Sequence 3, Appl
33	672.6	23.4	3943	2	US-08-852-091-3	Sequence 3, Appl
34	672.6	23.4	3943	2	US-08-820-754-3	Sequence 3, Appl
35	672.6	23.4	3943	3	US-08-956-652-3	Sequence 3, Appl
36	672.6	23.4	3943	3	US-08-956-869-3	Sequence 3, Appl
37	672.6	23.4	3943	3	US-08-948-547-3	Sequence 3, Appl
38	672.6	23.4	4003	5	PCT-US95-17025-3	Sequence 3, Appl
39	672.6	23.4	4003	4	US-09-087-465-1	Sequence 1, Appl
40	666.2	23.2	3943	4	US-09-364-970-10	Sequence 10, Appl
41	564.4	19.7	2375	1	US-08-369-796-9	Sequence 9, Appl
42	564.4	19.7	2375	2	US-08-852-091-9	Sequence 9, Appl
43	564.4	19.7	2375	2	US-08-820-754-9	Sequence 9, Appl
44	564.4	19.7	2375	3	US-08-956-652-9	Sequence 9, Appl
45	564.4	19.7	2375	3	US-08-956-869-9	Sequence 9, Appl

RESULT 1

US-08-369-796-11
Sequence 11, Application US/08369796
Patent No. 5716622
GENERAL INFORMATION:
APPLICANT: James E. Darnell, Jr.
APPLICANT: Zilong Wen
APPLICANT: Curt M. Horvath
APPLICANT: Zhong Zhong
TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/369,796
FILING DATE: 06-JAN-1995
CLASSIFICATION: A35
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET INFORMATION: 600-1-116
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 2869 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mouse
IMMEDIATE SOURCE:
LIBRARY: splenic/rhymic
CLONE: Murine 19s16

FEATURE:
NAME/KEY: CDS
LOCATION: 69..2378
US-08-369-796-11

Query Match 100.0%; Score 2869; DB 1; Length 2869;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCGACGACGAGCGCGCGCTCAGCGCGCTCAGCGCGAGACAGTGCAGACCCCTGACT 60
DB 1 GCGCGACGACGAGCGCGCGCTCAGCGCGCTCAGCGCGAGACAGTGCAGACCCCTGACT 60
QY 61 GCAGCAGATGCTCAGTGCAGACAGCTGCAGACAGTGCAGACAGCTTACTGAGAGC 120
DB 61 GCAGCAGATGCTCAGTGCAGACAGCTGCAGACAGTGCAGACAGCTTACTGAGAGC 120
QY 121 TGCACGAGTGTACAGGACGATTCCTCCATGAGCTGCGGAGATTCTGACCTTGA 180
DB 121 TGCACGAGTGTACAGGACGATTCCTCCATGAGCTGCGGAGATTCTGACCTTGA 180
QY 181 TTGAGAGTCAAGAGTGGCATATGACAGCAAGAGTCAATGCGATGCTGTTTC 240
DB 181 TTGAGAGTCAAGAGTGGCATATGACAGCAAGAGTCAATGCGATGCTGTTTC 240
QY 241 ATATATCTTGGGTGAATTTGACAGCAATATAGCCGATTCCTGCAAGAGTCCATGTC 300
DB 241 ATATATCTTGGGTGAATTTGACAGCAATATAGCCGATTCCTGCAAGAGTCCATGTC 300
QY 301 TCTATCAGCACAACCTTGAAGATCAAGAGTTCGACAGAGTATCTTGAAGC 360
DB 301 TCTATCAGCACAACCTTGAAGATCAAGAGTTCGACAGAGTATCTTGAAGC 360
QY 361 CAATGGAATTTGGCGGATGCTGCGCGATGCTGCGGAGAGATGCTGCTCTCCAGA 420
DB 361 CAATGGAATTTGGCGGATGCTGCGCGATGCTGCGGAGAGATGCTGCTCTCCAGA 420
QY 421 CGGAGCGACGCGAGCGCGACAGGAGGCGGACGCGCAACCGACCGCGCTAGTGA 480
DB 421 CGGAGCGACGCGAGCGCGACAGGAGGCGGACGCGCAACCGACCGCGCTAGTGA 480
QY 481 CAGAGAGCAGCAGATGTTGAGCAGCATCTTCAGGATGTCGGAACGAGTGCAGATC 540
DB 481 CAGAGAGCAGCAGATGTTGAGCAGCATCTTCAGGATGTCGGAACGAGTGCAGATC 540
QY 541 TAGAACGAAAATGAGGTGAGAGACCTCAGAGAGATTTGATTCACTACAAA 600
DB 541 TAGAACGAAAATGAGGTGAGAGACCTCAGAGAGATTTGATTCACTACAAA 600
QY 601 CCTTCAAGAGCGCAGAGACATGAGATGGAATGGAACACCAAGTCTGTGACAGAC 660
DB 601 CCTTCAAGAGCGCAGAGACATGAGATGGAATGGAACACCAAGTCTGTGACAGAC 660
QY 661 AGAAGATGACAGCAGCTGGAACAGATGCTCAGAGCGCTGAGAGAGGAGAGCAATTG 720
DB 661 AGAAGATGACAGCAGCTGGAACAGATGCTCAGAGCGCTGAGAGAGGAGAGCAATTG 720
QY 721 TGAATGAGCTGCGGGGCTCTTGTCAACAATGAGTACGTGAGAAACACTGACTATG 780
DB 721 TGAATGAGCTGCGGGGCTCTTGTCAACAATGAGTACGTGAGAAACACTGACTATG 780
QY 781 AAGAGCTGCGTGAAGAGGCGGCGAGAGATGCGTGCATGGAGCGCCCTCCAAACA 840
DB 781 AAGAGCTGCGTGAAGAGGCGGCGAGAGATGCGTGCATGGAGCGCCCTCCAAACA 840
QY 841 TCTGCTGAGACCTCTGGAAGAACTGATTAATCTTATGAGATTCATCTAGAGCC 900
DB 841 TCTGCTGAGACCTCTGGAAGAACTGATTAATCTTATGAGATTCATCTAGAGCC 900
QY 901 GCCAACAATTTAGAAATGAGAGAGCTGCAGCAGAAAGTCTCTACAGGCGACCTTA 960
DB 901 GCCAACAATTTAGAAATGAGAGAGCTGCAGCAGAAAGTCTCTACAGGCGACCTTA 960

QY 961 TCGTGCAGCACCAGCCCATGCTGAGAGAGACATCTGTGAGCTGTTCAGAACTTAATGA 1020
DB 961 TCGTGCAGCACCAGCCCATGCTGAGAGAGACATCTGTGAGCTGTTCAGAACTTAATGA 1020
QY 1021 AGAGTGCCTTCGTGTGAGAGCGGCGAGCGCTGCATGCGCATGACCGGAGCGGCTTAG 1080
DB 1021 AGAGTGCCTTCGTGTGAGAGCGGCGAGCGCTGCATGCGCATGACCGGAGCGGCTTAG 1080
QY 1081 TCATCAAGACTGTGTTCACATTTACACGAAAGTGTGCTGCTCAATTTCTGAGT 1140
DB 1081 TCATCAAGACTGTGTTCACATTTACACGAAAGTGTGCTGCTCAATTTCTGAGT 1140
QY 1141 TGAATTAATGCTTAAATTAAGTGTGATTAAGATCTGGGAGATGTTGCTGCC 1200
DB 1141 TGAATTAATGCTTAAATTAAGTGTGATTAAGATCTGGGAGATGTTGCTGCC 1200
QY 1201 TCAGAGGCTCTCGAAATTTAATTTACATTTCTGGGACGAAACACAAAGTATGACATGAGG 1260
DB 1201 TCAGAGGCTCTCGAAATTTAATTTACATTTCTGGGACGAAACACAAAGTATGACATGAGG 1260
QY 1261 AGTCTAACAGCGACGCTGTGAGAGTTCAGAGCTGACGCTTAAAGAGTATGAGAT 1320
DB 1261 AGTCTAACAGCGACGCTGTGAGAGTTCAGAGCTGACGCTTAAAGAGTATGAGAT 1320
QY 1321 GTGGAAATGAGGCGCGTGCATTTGATGCTGCTGATGCTGACTGAGAGCTGCAC 1380
DB 1321 GTGGAAATGAGGCGCGTGCATTTGATGCTGCTGATGCTGACTGAGAGCTGCAC 1380
QY 1381 TGATCAGCTTCGAGACTGAGCTGACCAAGAGCTCAAGATTTGACTTGAACCCACT 1440
DB 1381 TGATCAGCTTCGAGACTGAGCTGACCAAGAGCTCAAGATTTGACTTGAACCCACT 1440
QY 1441 CCTTGCAGTTGGGTGATGCTGCAACATGCTGATGCTGCAATGCTTGGCATCAATCC 1500
DB 1441 CCTTGCAGTTGGGTGATGCTGCAACATGCTGATGCTGCAATGCTTGGCATCAATCC 1500
QY 1501 TGTGTATACATGCTGACCAATTAACCCAGAGAGCTGAACTTCTTCACTTAAGCCGCAA 1560
DB 1501 TGTGTATACATGCTGACCAATTAACCCAGAGAGCTGAACTTCTTCACTTAAGCCGCAA 1560
QY 1561 TTGGAACCTGGGAGCAAGTGGCGGAGGTGCTGAGCTGCGCTGCTGCGACCAAGC 1620
DB 1561 TTGGAACCTGGGAGCAAGTGGCGGAGGTGCTGAGCTGCGCTGCTGCGACCAAGC 1620
QY 1621 GAGGCTTGAGCATGACAGCTGACCAAGCTGCTGAGAACTCTTGAAGGCTGTGTGA 1680
DB 1621 GAGGCTTGAGCATGACAGCTGACCAAGCTGCTGAGAACTCTTGAAGGCTGTGTGA 1680
QY 1681 ACTTACTAGGCTGTCAATATCATATGAGGCTAAATTTCTCAAGAAACATGCTGCGCAAGG 1740
DB 1681 ACTTACTAGGCTGTCAATATCATATGAGGCTAAATTTCTCAAGAAACATGCTGCGCAAGG 1740
QY 1741 GCTTCTCTTCTGGGCTGTGCTAGACAAATATCATGCACTTGTGAAATAATATCTTGG 1800
DB 1741 GCTTCTCTTCTGGGCTGTGCTAGACAAATATCATGCACTTGTGAAATAATATCTTGG 1800
QY 1801 CCTTTGGAATGAAGGCTAATCATGATGCTTCACTCAAGAGAGCGGAGCGGCGCATCC 1860
DB 1801 CCTTTGGAATGAAGGCTAATCATGATGCTTCACTCAAGAGAGCGGAGCGGCGCATCC 1860
QY 1861 TTAACACAAAGCGCGGAGCGCTTCTACTGGGCTTCAAGAGAGAGAGAGAGAGAG 1920
DB 1861 TTAACACAAAGCGCGGAGCGCTTCTACTGGGCTTCAAGAGAGAGAGAGAGAGAGAG 1920
QY 1921 GGTCTACTTTTCACTTGGGTGAAAGAGACATCAATGAGAGAGAGAGAGAGAGAGAG 1980
DB 1921 GGTCTACTTTTCACTTGGGTGAAAGAGACATCAATGAGAGAGAGAGAGAGAGAGAG 1980
QY 1981 AGCCATACAGCAGCAGCTGACAAACATGTCATTTGCTGAAATCATATGAGGCTATA 2040
DB 1981 AGCCATACAGCAGCAGCTGACAAACATGTCATTTGCTGAAATCATATGAGGCTATA 2040

Y	2041	AGATATGGATGTCGACACCATCCTGTGTCCTGACCTGTGCTACCTACCCGACATTC	2100
Db	2041	AGATATGGATGTCGACACCATCCTGTGTCCTGACCTGTGCTACCTACCCGACATTC	2100
Y	2101	CCAAAGGAGGAGGCAATTTGGAAAGTACTGTAGGCCGAGAGCCAGGAGCACCCCGAAGCG	2160
Db	2101	CCAAAGGAGGAGGCAATTTGGAAAGTACTGTAGGCCGAGAGCCAGGAGCACCCCGAAGCG	2160
Y	2161	ACCCAGGTAGTGTCTGCCCCGTACTCTGAAGACCAAGTTGATCTGTGTACACCAAGACT	2220
Db	2161	ACCCAGGTAGTGTCTGCCCCGTACTCTGAAGACCAAGTTGATCTGTGTACACCAAGACT	2220
Y	2221	GCAGCAATACCAATTGACCTGCGCATGTCCCGGACACTTATGATTCATTTGATGAGTTTG	2280
Db	2221	GCAGCAATACCAATTGACCTGCGCATGTCCCGGACACTTATGATTCATTTGATGAGTTTG	2280
Y	2281	GAAATTAACCGTAAAGGTGCTGACGCCCTCACAGAGGAGCGAGTTGAGTCCCTCACGTTTG	2340
Db	2281	GAAATTAACCGTAAAGGTGCTGACGCCCTCACAGAGGAGCGAGTTGAGTCCCTCACGTTTG	2340
Y	2341	ACATGATGTGACCTCGGAGTGTGTACCTCCGCCATGTGAGAGCTGGAACCAAGAACT	2400
Db	2341	ACATGATGTGACCTCGGAGTGTGTACCTCCGCCATGTGAGAGCTGGAACCAAGAACT	2400
Y	2401	GCAGAGACGTGACCTTGAGACACCTGCCCCCTGCTCCACCCCTAAGCAGCCGACCCCATTA	2460
Db	2401	GCAGAGACGTGACCTTGAGACACCTGCCCCCTGCTCCACCCCTAAGCAGCCGACCCCATTA	2460
Y	2461	TGCTGGAAGACCTCTACTTGTGTGTCGACGATTTTTTTTATTTATTTCCACTTGTGCT	2520
Db	2461	TGCTGGAAGACCTCTACTTGTGTGTCGACGATTTTTTTTATTTATTTCCACTTGTGCT	2520
Y	2521	ATCTTTGGGCAATCTGCGGCACTTTTAAAGAGAGAAATGAGTGTGCGGTGATAAAC	2580
Db	2521	ATCTTTGGGCAATCTGCGGCACTTTTAAAGAGAGAAATGAGTGTGCGGTGATAAAC	2580
Y	2581	TGTTATGTAAAGAGAGAGACCTCTGAGTCTGGGAGATGGGGCTAAGACAGAAAGGAGGC	2640
Db	2581	TGTTATGTAAAGAGAGAGACCTCTGAGTCTGGGAGATGGGGCTAAGACAGAAAGGAGGC	2640
Y	2641	AAAGGGAGACACCTCTGTCCTGCTGCCGCTGACCTCTTTTTCAGACAGCTGGGGGTTGG	2700
Db	2641	AAAGGGAGACACCTCTGTCCTGCTGCCGCTGACCTCTTTTTCAGACAGCTGGGGGTTGG	2700
Y	2701	TTGTTAGACAAGTGCCTCTGCTGTCGCCATGAGCTACTGTTGCCACTCTGTGAGCTGAT	2760
Db	2701	TTGTTAGACAAGTGCCTCTGCTGTCGCCATGAGCTACTGTTGCCACTCTGTGAGCTGAT	2760
Y	2761	ACCCCTCTGCGGACCTCTGCGCTGACCTTTAACTGCTAATATCCACATGAAGC	2820
Db	2761	ACCCCTCTGCGGACCTCTGCGCTGACCTTTAACTGCTAATATCCACATGAAGC	2820
Y	2821	TAGGACTAAGCCCGAGAGGTTCTTTTAAATTAATAAAAAAAAAAAAAA	2869
Db	2821	TAGGACTAAGCCCGAGAGGTTCTTTTAAATTAATAATAAAAAAAAAAAAAA	2869

```

STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852,091
FILING DATE: 06-MAY-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/369,796
FILING DATE: 06-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-116
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 2869 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mouse
IMMEDIATE SOURCE:
LIBRARY: splenic/thymic
CLONE: Murine 19sf6
FEATURE:
NAME/KEY: CDS
LOCATION: 69..2378
US-08-852-091-11

Query Match 100.0%; Score 2869; DB 2; Length 2869;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GCGCGACAGCCAGCCGCGCCAGTCGGGCTCAGCCCGAGACAGTCGAGACCCCTGACT 60
Db 1 GCGCGACAGCCAGCCGCGCGCCAGTCGGGCTCAGCCCGAGACAGTCGAGACCCCTGACT 60
61 GCAGCAGATGGCTCAGTGAAGCAACAGCTGCAGCAGCTGGACACACCGTACCTGAAGCAGC 120
Qy 61 GCAGCAGATGGCTCAGTGAAGCAACAGCTGCAGCAGCTGGACACACCGTACCTGAAGCAGC 120
Db 61 GCAGCAGATGGCTCAGTGAAGCAACAGCTGCAGCAGCTGGACACACCGTACCTGAAGCAGC 120
121 TGCACACAGGTGACAGCGCAGCAGCTTCCCATGGAGGCTGGGAGTTCCTGGGACCTTGA 180
Qy 121 TGCACACAGGTGACAGCGCAGCAGCTTCCCATGGAGGCTGGGAGTTCCTGGGACCTTGA 180
Db 121 TGCACACAGGTGACAGCGCAGCAGCTTCCCATGGAGGCTGGGAGTTCCTGGGACCTTGA 180
181 TTGAGAGTCAAGACTGGGCATATGACAGCCAGCAAGAGAGTCACATGGCCAGCTGGTGTTC 240
Qy 181 TTGAGAGTCAAGACTGGGCATATGACAGCCAGCAAGAGAGTCACATGGCCAGCTGGTGTTC 240
Db 181 TTGAGAGTCAAGACTGGGCATATGACAGCCAGCAAGAGAGTCACATGGCCAGCTGGTGTTC 240
241 ATAACTCTCTGGGGTAAATTTGACAGCAATATAGCCGATTCTCTGCAAGAGTCCAAATGTC 300
Qy 241 ATAACTCTCTGGGGTAAATTTGACAGCAATATAGCCGATTCTCTGCAAGAGTCCAAATGTC 300
Db 241 ATAACTCTCTGGGGTAAATTTGACAGCAATATAGCCGATTCTCTGCAAGAGTCCAAATGTC 300
301 TCTATCAGACACACTTGAAGAATCAAGACAGATTCTCGAGAGCAGGTATCTTGAAGAC 360
Qy 301 TCTATCAGACACACTTGAAGAATCAAGACAGATTCTCGAGAGCAGGTATCTTGAAGAC 360
Db 301 TCTATCAGACACACTTGAAGAATCAAGACAGATTCTCGAGAGCAGGTATCTTGAAGAC 360
361 CATGGAATTGCCCGATCGTGGCCGATGCTGTGGAGAGTCTGCGCTCTCCAGA 420
Qy 361 CATGGAATTGCCCGATCGTGGCCGATGCTGTGGAGAGTCTGCGCTCTCCAGA 420

```

Db 361 CATGGAATTTGCCGATCTGTGCCCCGATGGCTGTGGGAAGATCTCGCTCTCCAGA 420
Oy 421 CGGACCCACGGCAGCCACGAAAGGCGCCAGGCCAACCCACAGCCGCCGTAGTGA 480
Db 421 CGGACCCACGGCAGCCACGAAAGGCGCCAGGCCAACCCACAGCCGCCGTAGTGA 480
Oy 481 CAGAGAAGCAGATGTTGGACGACATCTTGAAGATGTCGGGAAGCGAGTGCAGATG 540
Db 481 CAGAGAAGCAGATGTTGGACGACATCTTGAAGATGTCGGGAAGCGAGTGCAGATG 540
Oy 541 TAGAACGAAATTAAGATGTGTGGAGAACCTCCAGAGCAGATTTGATTTTCACTACAAAA 600
Db 541 TAGAACGAAATTAAGATGTGTGGAGAACCTCCAGAGCAGATTTGATTTTCACTACAAAA 600
Oy 601 CCTCAAGAGCCCAAGAGACATGACAGATCTGTAATGAAACCAACAGCTCTGTACCCAGAC 660
Db 601 CCTCAAGAGCCCAAGAGACATGACAGATCTGTAATGAAACCAACAGCTCTGTACCCAGAC 660
Oy 661 AGAAGATGCGAGCTGGAACAGATGCTACAGCCCTGAGACAGATGCGGAGAAGCATTTG 720
Db 661 AGAAGATGCGAGCTGGAACAGATGCTACAGCCCTGAGACAGATGCGGAGAAGCATTTG 720
Oy 721 TGAATGAGCTGGGGGGGCTCTTGTGACCAATGGAGTACGTGCAGAAAGACACTGATGATG 780
Db 721 TGAATGAGCTGGGGGGGCTCTTGTGACCAATGGAGTACGTGCAGAAAGACACTGATGATG 780
Oy 781 AAGAGCTGGCTGACTGGAAGAGCGCGCCAGAGATTCGGCTGCATCGGAGGCGCTCCACACA 840
Db 781 AAGAGCTGGCTGACTGGAAGAGCGCGCCAGAGATTCGGCTGCATCGGAGGCGCTCCACACA 840
Oy 841 TGTGCTGAGCCCTGTGGAAGAACTGATTAATCTTATGAGAGATTCATCACTTCAGACC 900
Db 841 TGTGCTGAGCCCTGTGGAAGAACTGATTAATCTTATGAGAGATTCATCACTTCAGACC 900
Oy 901 GCCAACAAATTAAGAACTGAGAGAGCTGCAGCAGAAAGTGTCTTACAGAGCGGACCTTA 960
Db 901 GCCAACAAATTAAGAACTGAGAGAGCTGCAGCAGAAAGTGTCTTACAGAGCGGACCTTA 960
Oy 961 TCGTGCGAGCAGCCGCGCCATGCTGGAGAGAGATCGTGGAGCTGTGAGAACTTAATGA 1020
Db 961 TCGTGCGAGCAGCCGCGCCATGCTGGAGAGAGATCGTGGAGCTGTGAGAACTTAATGA 1020
Oy 1021 AGATGCTCTGTGTGTGGAGCGGAGCCCTGTGATGCCATGCAACCCGAGCGGCCCTTAG 1080
Db 1021 AGATGCTCTGTGTGTGGAGCGGAGCCCTGTGATGCCATGCAACCCGAGCGGCCCTTAG 1080
Oy 1081 TCATCAAGACTGGTGTGCACTTTACACGAAAGTCAAGTGTGCTGTAAATTTCTGAGT 1140
Db 1081 TCATCAAGACTGGTGTGCACTTTACACGAAAGTCAAGTGTGCTGTAAATTTCTGAGT 1140
Oy 1141 TGAATTTCAAGCTTAAATTAAGTGTGATTTGATTAAGACTCTGGGGATGTTCTGCCCC 1200
Db 1141 TGAATTTCAAGCTTAAATTAAGTGTGATTTGATTAAGACTCTGGGGATGTTCTGCCCC 1200
Oy 1201 TCAGAGGGTCTCGGAATTTAACAATTTCTGGGACAGAAACAAAAAGTATGAATGAGG 1260
Db 1201 TCAGAGGGTCTCGGAATTTAACAATTTCTGGGACAGAAACAAAAAGTATGAATGAGG 1260
Oy 1261 AGCTTAACCAAGCGAGCTGTCTGACAGATTCAGACACCTTGAAGGAGCAGAGAT 1320
Db 1261 AGCTTAACCAAGCGAGCTGTCTGACAGATTCAGACACCTTGAAGGAGCAGAGAT 1320
Oy 1321 GTGGGATGAGGCGCGTGGCAATTTGATGCTCTTGTATGCTGATGAGAGAGCTGCACC 1380
Db 1321 GTGGGATGAGGCGCGTGGCAATTTGATGCTCTTGTATGCTGATGAGAGAGCTGCACC 1380
Oy 1381 TGATCACCTTCGAGACTGAGGTGTACCAAGCCCTCAAGATTTGACCTAGAGACCCACT 1440
Db 1381 TGATCACCTTCGAGACTGAGGTGTACCAAGCCCTCAAGATTTGACCTAGAGACCCACT 1440
Oy 1441 CCTTGGCAGTTGTGTATCTCCAAACATCTGTGATGATGCCAAATGCTTGGGATCAATCC 1500
Db 1441 CCTTGGCAGTTGTGTATCTCCAAACATCTGTGATGATGCCAAATGCTTGGGATCAATCC 1500

Oy 1501 TGTGTATAACATGCTGACCAATTAACCCCAAGAACGTGAACCTTCTACTAAGCCGCAA 1560
Db 1501 TGTGTATAACATGCTGACCAATTAACCCCAAGAACGTGAACCTTCTACTAAGCCGCAA 1560
Oy 1561 TTGGAACTGTGGACCAAGTGGCCGAGGTGCTCAGCTGGAGTTCTGTCCAGCAGAAC 1620
Db 1561 TTGGAACTGTGGACCAAGTGGCCGAGGTGCTCAGCTGGAGTTCTGTCCAGCAGAAC 1620
Oy 1621 GAGGCTGACCATCGAGCAGCAGCTGACAAACGCTGGCTGAGAAAGTCTTAAGGCCCTGTCTGA 1680
Db 1621 GAGGCTGACCATCGAGCAGCAGCTGACAAACGCTGGCTGAGAAAGTCTTAAGGCCCTGTCTGA 1680
Oy 1681 ACTACTGAGGTGTGACATGACATGAGGCTAAATTTCTCAAGAAACATGCTGTGCAAGG 1740
Db 1681 ACTACTGAGGTGTGACATGACATGAGGCTAAATTTCTCAAGAAACATGCTGTGCAAGG 1740
Oy 1741 GCTTCTCTCTGTGGTGTGCTAGACAAATATCATGACCTTGTGAAAGATATATCTTG 1800
Db 1741 GCTTCTCTCTGTGGTGTGCTAGACAAATATCATGACCTTGTGAAAGATATATCTTG 1800
Oy 1801 CCTTTGGAATGAAGGTTACATCATGAGTTTTCATCAGCAAGAGCCGAGCGGCATCC 1860
Db 1801 CCTTTGGAATGAAGGTTACATCATGAGTTTTCATCAGCAAGAGCCGAGCGGCATCC 1860
Oy 1861 TAAACAAAGCCCGCGGACCTTCTACTGCGCTTCAGCAGAGAGCAGCAAGAAAGAG 1920
Db 1861 TAAACAAAGCCCGCGGACCTTCTACTGCGCTTCAGCAGAGAGCAGCAAGAAAGAG 1920
Oy 1921 GGGTACCTTCTACTTGGGTGAAAGAGATCACTGTGCAAGCCAGATCCAGTCTGTAG 1980
Db 1921 GGGTACCTTCTACTTGGGTGAAAGAGATCACTGTGCAAGCCAGATCCAGTCTGTAG 1980
Oy 1981 AGCCATTAACCAAGCAGCAGCTGACAAACATGTTGCTGAATATCATGAGGCTATA 2040
Db 1981 AGCCATTAACCAAGCAGCAGCTGACAAACATGTTGCTGAATATCATGAGGCTATA 2040
Oy 2041 AGATCATGATGCGACCAACATCTGCTGTCCACTTGTCTACTTACCCGACATTC 2100
Db 2041 AGATCATGATGCGACCAACATCTGCTGTCCACTTGTCTACTTACCCGACATTC 2100
Oy 2101 CCAAGAGAGGCAATTTGAAAGTACTGTAGGCCCGAGCCAGAGCAGACCCGAAACCG 2160
Db 2101 CCAAGAGAGGCAATTTGAAAGTACTGTAGGCCCGAGAGCCAGAGCAGACCCGAAACCG 2160
Oy 2161 ACCGAGTACTGTGCTGCCCCGCTACTGAAAGACCAAGTTCATCTGTGACACCAAGCACT 2220
Db 2161 ACCGAGTACTGTGCTGCCCCGCTACTGAAAGACCAAGTTCATCTGTGACACCAAGCACT 2220
Oy 2221 GCAGCAATACCATGACCTGCGGATGTCGCCCGCAGCTTATGATTCATGATGAGCTTTG 2280
Db 2221 GCAGCAATACCATGACCTGCGGATGTCGCCCGCAGCTTATGATTCATGATGAGCTTTG 2280
Oy 2281 GAAATTAACGTTGAAGGTGTGAGCCCTCAGCAGAGGCGAGTTGAGTCCCTACGTTTG 2340
Db 2281 GAAATTAACGTTGAAGGTGTGAGCCCTCAGCAGAGGCGAGTTGAGTCCCTACGTTTG 2340
Oy 2341 ACATGATCTGACCTGCGAGTGTCTACCTCCCATGTGAGAGGCTGAAACCAAGAGCT 2400
Db 2341 ACATGATCTGACCTGCGAGTGTCTACCTCCCATGTGAGAGGCTGAAACCAAGAGCT 2400
Oy 2401 GCAGAGACGTGACTTGAAGACACCTGCCCCGCTGCACACCCCTAAGCAGCCGAAACCCATA 2460
Db 2401 GCAGAGACGTGACTTGAAGACACCTGCCCCGCTGCACACCCCTAAGCAGCCGAAACCCATA 2460
Oy 2461 TCGTCTGAACCTCTAATCTTGTGTGCTCAGATTTTCTTAAATTTCTACTTGTCT 2520
Db 2461 TCGTCTGAACCTCTAATCTTGTGTGCTCAGATTTTCTTAAATTTCTACTTGTCT 2520
Oy 2521 ATCTTTGGCAATCTGGCACTTTTAAAGAGGAATAGTAGTGGGTGATTAAC 2580
Db 2521 ATCTTTGGCAATCTGGCACTTTTAAAGAGGAATAGTAGTGGGTGATTAAC 2580

Dh 721 TGAAGTACGCTGGCGGCTCTTGTACCAATGAGATCTGACAGACACTGACTGATG 780
Oy 781 AAGAGCTGGCTGACTGGAGAGCGCGCAGAGATCGCTGATGGAGGCCCTCCACACA 840
Dh 781 AAGAGCTGGCTGACTGGAGAGCGCGCAGAGATCGCTGATGGAGGCCCTCCACACA 840
Oy 841 TCTGCTGGAGCGCTGTGAAAACTGATTAATTCATTAGCAGAACTCTCACTTCAGACCC 900
Dh 841 TCTGCTGGAGCGCTGTGAAAACTGATTAATTCATTAGCAGAACTCTCACTTCAGACCC 900
Oy 901 GCCAACAAATTAAGAACTGGAGAGCTGACGCAAGAAAGTTCCTACAGAGGCCACCCCTA 960
Dh 901 GCCAACAAATTAAGAACTGGAGAGCTGACGCAAGAAAGTTCCTACAGAGGCCACCCCTA 960
Oy 961 TCGTGCAGACACCGCCCATGCTGGAGAGAGATCGTGGAGCTGTTGCAAACTTAATGA 1020
Dh 961 TCGTGCAGACACCGCCCATGCTGGAGAGAGATCGTGGAGCTGTTGCAAACTTAATGA 1020
Oy 1021 AGAGTCCCTTCTGTGTGGAGCGGAGCCCTGATGCCATGCCACCGGAGCCGCTTAG 1080
Dh 1021 AGAGTCCCTTCTGTGTGGAGCGGAGCCCTGATGCCATGCCACCGGAGCCGCTTAG 1080
Oy 1081 TCATCAAGACGCTGTGAGTTACACAGAAAGTCAGGTTGCTGGTCAATTTCTGAGT 1140
Dh 1081 TCATCAAGACGCTGTGAGTTACACAGAAAGTCAGGTTGCTGGTCAATTTCTGAGT 1140
Oy 1141 TGAATTAACAGCTTAATAATTAAGTGTGATTAAGACCTGGGGATTTGCTGCC 1200
Dh 1141 TGAATTAACAGCTTAATAATTAAGTGTGATTAAGACCTGGGGATTTGCTGCC 1200
Oy 1201 TCAGAGGCTCTCGGAAATTTAACATTTCTGGCAGAGACACAAAAGTATGAACATGAGG 1260
Dh 1201 TCAGAGGCTCTCGGAAATTTAACATTTCTGGCAGAGACACAAAAGTATGAACATGAGG 1260
Oy 1261 AGTAAACAGCGAGCGCTGCTGACAGATTCAAGACCTGACCCCTTGGGAGCGAGAT 1320
Dh 1261 AGTAAACAGCGAGCGCTGCTGACAGATTCAAGACCTGACCCCTTGGGAGCGAGAT 1320
Oy 1321 GTGGAAATGAGCGCGTGGCCAAATTTGTATGCTCTTGTATGCTGATGAGGAGCTGCAC 1380
Dh 1321 GTGGAAATGAGCGCGTGGCCAAATTTGTATGCTCTTGTATGCTGATGAGGAGCTGCAC 1380
Oy 1381 TGATCACCCTTGAGAGTGAAGGTATACCAAGCGCTCAAGTTCAGTACAGACCCCT 1440
Dh 1381 TGATCACCCTTGAGAGTGAAGGTATACCAAGCGCTCAAGTTCAGTACAGACCCCT 1440
Oy 1441 CCTTGCAGATTGTGATCTCAACAATCTGTCAAGATGCCAAATGCTTGGGCATCAATCC 1500
Dh 1441 CCTTGCAGATTGTGATCTCAACAATCTGTCAAGATGCCAAATGCTTGGGCATCAATCC 1500
Oy 1501 TGTGCTATTAACATGCTGACCAATTAACCCCAAGACGTGAACCTTCTCAAGCGCCAA 1560
Dh 1501 TGTGCTATTAACATGCTGACCAATTAACCCCAAGACGTGAACCTTCTCAAGCGCCAA 1560
Oy 1561 TTGGAACCTGGAGACCAAGTGGCCGAGGTCTGACGTGCAATTTCTGCTCCACCAAGC 1620
Dh 1561 TTGGAACCTGGAGACCAAGTGGCCGAGGTCTGACGTGCAATTTCTGCTCCACCAAGC 1620
Oy 1621 GAGGCTGAGACATCGACAGCTGACAAAGCGCTGAGAAAGCTCTAGGGCTGTGTA 1680
Dh 1621 GAGGCTGAGACATCGACAGCTGACAAAGCGCTGAGAAAGCTCTAGGGCTGTGTA 1680
Oy 1681 ACTAGTGAAGGTGTACATGACATGGGCTTAATCTGAAAGAAACATGGCTGGCAAGG 1740
Dh 1681 ACTAGTGAAGGTGTACATGACATGGGCTTAATCTGAAAGAAACATGGCTGGCAAGG 1740
Oy 1741 GCTTCTCTCTGTGGTCTGCTAGACATATCATGACCTGTGAAAAATATATCTTGG 1800
Dh 1741 GCTTCTCTCTGTGGTCTGCTAGACATATCATGACCTGTGAAAAATATATCTTGG 1800
Oy 1801 CCTTTTGAATGAAGGTACATCATGGGTTTCATCAGCAAGAGCGGAGCGGCATCC 1860
Dh 1801 CCTTTTGAATGAAGGTACATCATGGGTTTCATCAGCAAGAGCGGAGCGGCATCC 1860

Oy 1861 TAAGCACAAGCCCCGGGACCTTCTACTGCGCTTCAGCGAGAGCAGCAAAAGAGAG 1920
Dh 1861 TAAGCACAAGCCCCGGGACCTTCTACTGCGCTTCAGCGAGAGCAGCAAAAGAGAG 1920
Oy 1921 GGGTACTTTTCACTTGGGTGAAAAAGACATCAGTGGCAGACCCAGATCCAGTGTAG 1980
Dh 1921 GGGTACTTTTCACTTGGGTGAAAAAGACATCAGTGGCAGACCCAGATCCAGTGTAG 1980
Oy 1981 AGCCATACCCAGAGAGCTGAACCAACATGTCATTTGCTGAATTCATCATGGCTTGA 2040
Dh 1981 AGCCATACCCAGAGAGCTGAACCAACATGTCATTTGCTGAATTCATCATGGCTTGA 2040
Oy 2041 AGATCATGGATGCGACCAACATCTGTGTGTCACATTTGTACCTTAACCCGACATTC 2100
Dh 2041 AGATCATGGATGCGACCAACATCTGTGTGTCACATTTGTACCTTAACCCGACATTC 2100
Oy 2101 CCAGGAGAGGCAATTTGAAAGTACTGTAGGCCGAGAGCCAGAGACCCCGAAGCG 2160
Dh 2101 CCAGGAGAGGCAATTTGAAAGTACTGTAGGCCGAGAGCCAGAGACCCCGAAGCG 2160
Oy 2161 ACCCAGTAGTGTGCGCCCGTACCTGAAGACCAATTCATGCTGTGACACCAAGACT 2220
Dh 2161 ACCCAGTAGTGTGCGCCCGTACCTGAAGACCAATTCATGCTGTGACACCAAGACT 2220
Oy 2221 GCAGCAATACCAATTAACCTGCGATGTCGCCCGACCTTAATTCATTTGATGACGTTTG 2280
Dh 2221 GCAGCAATACCAATTAACCTGCGATGTCGCCCGACCTTAATTCATTTGATGACGTTTG 2280
Oy 2281 GAATTAAGCGTGAAGGTGCTGACCCCTGACGAGAGGCGAGTTTGAAGTCTGACGTTTG 2340
Dh 2281 GAATTAAGCGTGAAGGTGCTGACCCCTGACGAGGAGGCGAGTTTGAAGTCTGACGTTTG 2340
Oy 2341 ACATGATCTGACCTCGGAGTGTGCTACCTCCCATGAGAGAGCTGAACACAGAGCT 2400
Dh 2341 ACATGATCTGACCTCGGAGTGTGCTACCTCCCATGAGAGAGCTGAACACAGAGCT 2400
Oy 2401 GCAGAGAGTGACTTGAAGACCTGCCCCGTGCTCCACCCCTAAGCAGCCGAACCCATA 2460
Dh 2401 GCAGAGAGTGACTTGAAGACCTGCCCCGTGCTCCACCCCTAAGCAGCCGAACCCATA 2460
Oy 2461 TCGCTGAAACCTCAACTTTGTGTGCTCCAGATTTTAAAAATTTCTCACTTCTGCT 2520
Dh 2461 TCGCTGAAACCTCAACTTTGTGTGCTCCAGATTTTAAAAATTTCTCACTTCTGCT 2520
Oy 2521 ATCTTTGGGCAATCTGGGCACTTTTAAAGAGAAATGAGTGAAGTGGGTGAATAAC 2580
Dh 2521 ATCTTTGGGCAATCTGGGCACTTTTAAAGAGAAATGAGTGAAGTGGGTGAATAAC 2580
Oy 2581 TGTATGTAAAGAGAGAGACCTCTGAGTCTGGGATGGGCTGAGACAGAAAGGAGGC 2640
Dh 2581 TGTATGTAAAGAGAGAGACCTCTGAGTCTGGGATGGGCTGAGACAGAAAGGAGGC 2640
Oy 2641 AAAGGGAAACCTCTGTCGCGCGCGCTGCGCTCTTTTACAGACCTGGGGGTGG 2700
Dh 2641 AAAGGGAAACCTCTGTCGCGCGCGCTGCGCTCTTTTACAGACCTGGGGGTGG 2700
Oy 2701 TTGTTAGACAAGTGCCTCTGTGCGCATGGCTACCTTGTGCCCACTGTGAGCTGAT 2760
Dh 2701 TTGTTAGACAAGTGCCTCTGTGCGCATGGCTACCTTGTGCCCACTGTGAGCTGAT 2760
Oy 2761 ACCCATCTGGGAACTCTGTGCTGTGACATTTCAACTTGTCTAATATTCACATAGAAGC 2820
Dh 2761 ACCCATCTGGGAACTCTGTGCTGTGACATTTCAACTTGTCTAATATTCACATAGAAGC 2820
Oy 2821 TAGGACTAAGCCAGAGAGTCTCTTTAATTTAAAAAATTTAAAAA 2869
Dh 2821 TAGGACTAAGCCAGAGAGTCTCTTTAATTTAAAAAATTTAAAAA 2869

RESULT 4
US-08-956-652-11
: Sequence 11, Application US/08956652

Patent No. 6013475
GENERAL INFORMATION:
APPLICANT: Darnell Jr., James E.
APPLICANT: Schindler, Christian W.
APPLICANT: Fu, Xian-Yuan
APPLICANT: Men, Zilong
APPLICANT: Zhong, Zhong
TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klaber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,652
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,185
FILING DATE: 11-MAR-1994
APPLICATION NUMBER: US 07/980,498
FILING DATE: 23-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/854,296
FILING DATE: 19-MAR-1992
APPLICATION NUMBER: WO US93/02569
FILING DATE: 19-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,588
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-073 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 2869 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mouse
IMMEDIATE SOURCE:
LIBRARY: splenic/lymfc
CLONE: Murine 19s16
FEATURE:
NAME/KEY: CDS
LOCATION: 69..2378
US-08-956-652-11

Query Match 100.0%; Score 2869; DB 3; Length 2869;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCGCGACAGCCAGCGCGCGCTCAGCGCCGAGACAGCTGAGACCCCTGACT 60

|||||
Db 1 GCCGCGACAGCCAGCGCGCGCTCAGCGCCGAGACAGCTGAGACCCCTGACT 60
QY 61 GCAGCAGATGCGCTGACTGGAACAGCTGCAGCAGCTGCAACACGCTACTGAAAGC 120
|||||
Db 61 GCAGCAGATGCGCTGACTGGAACAGCTGCAGCAGCTGCAACACGCTACTGAAAGC 120
QY 121 TGCACAGCTGTACAGGACACGTTCCCTATGAGCTGCGGAGCTTCTTGACCTTGA 180
121 TGCACAGCTGTACAGGACACGTTCCCTATGAGCTGCGGAGCTTCTTGACCTTGA 180
Db 121 TGCACAGCTGTACAGGACACGTTCCCTATGAGCTGCGGAGCTTCTTGACCTTGA 180
QY 181 TTGAGAGTCAAGAGCTGGCATATGACCCAGCAAGAGTACATGCCAGCTTGTGTTTC 240
181 TTGAGAGTCAAGAGCTGGCATATGACCCAGCAAGAGTACATGCCAGCTTGTGTTTC 240
Db 181 TTGAGAGTCAAGAGCTGGCATATGACCCAGCAAGAGTACATGCCAGCTTGTGTTTC 240
QY 241 ATATCTCTTGGGGAATGACACGAATATACCCGATTCCTCCAGAGAGTCAATGTC 300
241 ATATCTCTTGGGGAATGACACGAATATACCCGATTCCTCCAGAGAGTCAATGTC 300
Db 241 ATATCTCTTGGGGAATGACACGAATATACCCGATTCCTCCAGAGAGTCAATGTC 300
QY 301 TCTATCAGCAACACTTTCGAAGATCAAGCAGTTTCTGCAGCAGGTATCTGGAAGC 360
301 TCTATCAGCAACACTTTCGAAGATCAAGCAGTTTCTGCAGCAGGTATCTGGAAGC 360
Db 301 TCTATCAGCAACACTTTCGAAGATCAAGCAGTTTCTGCAGCAGGTATCTGGAAGC 360
QY 361 CAATGGAATTTGCCGATGCTGCGCGATGCTCTGTGGGAAGATCTCGCTCTTACA 420
361 CAATGGAATTTGCCGATGCTGCGCGATGCTCTGTGGGAAGATCTCGCTCTTACA 420
Db 361 CAATGGAATTTGCCGATGCTGCGCGATGCTCTGTGGGAAGATCTCGCTCTTACA 420
QY 421 CGCAGCAGCGCAGCGCCAGCGAGCGGCGCAGCGCAACCGCAAGCGCGCTACTGA 480
421 CGCAGCAGCGCAGCGCCAGCGAGCGGCGCAGCGCAACCGCAAGCGCGCTACTGA 480
Db 421 CGCAGCAGCGCAGCGCCAGCGAGCGGCGCAGCGCAACCGCAAGCGCGCTACTGA 480
QY 481 CAGAGAAGCAGCAGATGTTGAGCAGCATCTTCAGATGTCGCGAGCAGTGCAGATC 540
481 CAGAGAAGCAGCAGATGTTGAGCAGCATCTTCAGATGTCGCGAGCAGTGCAGATC 540
Db 481 CAGAGAAGCAGCAGATGTTGAGCAGCATCTTCAGATGTCGCGAGCAGTGCAGATC 540
QY 541 TAGAACGAAATGAGAGTGTGAGAACCTCCAGACACTTTGATTTACATACAAA 600
541 TAGAACGAAATGAGAGTGTGAGAACCTCCAGACACTTTGATTTACATACAAA 600
Db 541 TAGAACGAAATGAGAGTGTGAGAACCTCCAGACACTTTGATTTACATACAAA 600
QY 601 CCTCAAGACCCAGGACATGACAGATCTGAATGGAACCAACCACTTGTGACAGAC 660
601 CCTCAAGACCCAGGACATGACAGATCTGAATGGAACCAACCACTTGTGACAGAC 660
Db 601 CCTCAAGACCCAGGACATGACAGATCTGAATGGAACCAACCACTTGTGACAGAC 660
QY 661 AGAAGATGACAGAGCTGGAACAGATGCTCACAGCCTTGAGCAGAGTGCAGGAATG 720
661 AGAAGATGACAGAGCTGGAACAGATGCTCACAGCCTTGAGCAGAGTGCAGGAATG 720
Db 661 AGAAGATGACAGAGCTGGAACAGATGCTCACAGCCTTGAGCAGAGTGCAGGAATG 720
QY 721 TGAGTGAAGTGGGGGGCTTGTCAAGCAATGGAGTACGTCAGAAACACTGATG 780
721 TGAGTGAAGTGGGGGGCTTGTCAAGCAATGGAGTACGTCAGAAACACTGATG 780
Db 721 TGAGTGAAGTGGGGGGCTTGTCAAGCAATGGAGTACGTCAGAAACACTGATG 780
QY 781 AAGAGTGCCTGACTGGAAGAGCGGCGCAGAGATCGGTCATGCGGCGCTGCCACA 840
781 AAGAGTGCCTGACTGGAAGAGCGGCGCAGAGATCGGTCATGCGGCGCTGCCACA 840
Db 781 AAGAGTGCCTGACTGGAAGAGCGGCGCAGAGATCGGTCATGCGGCGCTGCCACA 840
QY 841 TCTGCTGACCGCTGTGAAACCTGATACTTATGACGATCTCAACTTCAGACC 900
841 TCTGCTGACCGCTGTGAAACCTGATACTTATGACGATCTCAACTTCAGACC 900
Db 841 TCTGCTGACCGCTGTGAAACCTGATACTTATGACGATCTCAACTTCAGACC 900
QY 901 GCCAACAAATTAAGAACTGAGAGCTGCAGAGAAAGTCTCTACAGGGGACCTTA 960
901 GCCAACAAATTAAGAACTGAGAGCTGCAGAGAAAGTCTCTACAGGGGACCTTA 960
Db 901 GCCAACAAATTAAGAACTGAGAGCTGCAGAGAAAGTCTCTACAGGGGACCTTA 960
QY 961 TCGTGCAGCAGCGGCCCATGCTGAGAGAGAGATCGTGGAGCTGTTCAAGAACTTAATGA 1020
961 TCGTGCAGCAGCGGCCCATGCTGAGAGAGAGATCGTGGAGCTGTTCAAGAACTTAATGA 1020
Db 961 TCGTGCAGCAGCGGCCCATGCTGAGAGAGAGATCGTGGAGCTGTTCAAGAACTTAATGA 1020
QY 1021 AGAGTCCCTTCTGATGAGAGCGGCGCAGCTGATGCCATGCAACCGGACCGGCTTAG 1080
1021 AGAGTCCCTTCTGATGAGAGCGGCGCAGCTGATGCCATGCAACCGGACCGGCTTAG 1080
Db 1021 AGAGTCCCTTCTGATGAGAGCGGCGCAGCTGATGCCATGCAACCGGACCGGCTTAG 1080
QY 1081 TCATCAAGAGCTGTGTCAGTTTACCAAGCAAGTCAAGTTGCTGTGTAATTTCTGAGT 1140
1081 TCATCAAGAGCTGTGTCAGTTTACCAAGCAAGTCAAGTTGCTGTGTAATTTCTGAGT 1140
|||||

Db 1081 TCATCAAGACTGCTGTCAGCTTTTACACGAAGTCAAGTCTTCTGCTCAAAATTTCTGAGT 1140
Qy 1141 TGAATTAATCAGCTTTAAATTAATTAAGTGTGATTAAGAGCTGTGGGATGTTCTGCCCC 1200
Db 1141 TGAATTAATCAGCTTTAAATTAATTAAGTGTGATTAAGAGCTGTGGGATGTTCTGCCCC 1200
Qy 1201 TCGAGGCTCTCGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260
Db 1201 TCGAGGCTCTCGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260
Qy 1261 AGTCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
Db 1261 AGTCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
Qy 1321 GTGGGAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
Db 1321 GTGGGAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
Qy 1381 TGATCAGCTTCGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 1440
Db 1381 TGATCAGCTTCGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 1440
Qy 1441 CCTTGCAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1500
Db 1441 CCTTGCAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1500
Qy 1501 TGTGCTTAATCAATGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1560
Db 1501 TGTGCTTAATCAATGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1560
Qy 1561 TTGGGAACCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1620
Db 1561 TTGGGAACCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1620
Qy 1621 GAGGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1680
Db 1621 GAGGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1680
Qy 1681 ACTACTAGGAGTGTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1740
Db 1681 ACTACTAGGAGTGTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1740
Qy 1741 GCTTCT 1800
Db 1741 GCTTCT 1800
Qy 1801 CCTTTCAGATGAAGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Db 1801 CCTTTCAGATGAAGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Qy 1861 TAAACACAAAGCCCGGAGGAGCTTCT 1920
Db 1861 TAAACACAAAGCCCGGAGGAGCTTCT 1920
Qy 1921 GGGTCACTTCT 1980
Db 1921 GGGTCACTTCT 1980
Qy 1981 AGCCATACACCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2040
Db 1981 AGCCATACACCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2040
Qy 2041 AGATCATGATGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2100
Db 2041 AGATCATGATGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2100
Qy 2101 CCAAGGAG 2160
Db 2101 CCAAGGAG 2160
Qy 2161 ACCAGGAG 2220
Db 2161 ACCAGGAG 2220

Qy 2221 GCACCAATACATTTGACCTGCGCCATGTCGCCCGCAGCTTTAGATTCAATGATGAGTTTG 2280
Db 2221 GCACCAATACATTTGACCTGCGCCATGTCGCCCGCAGCTTTAGATTCAATGATGAGTTTG 2280
Qy 2281 GAAATTAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2340
Db 2281 GAAATTAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2340
Qy 2341 ACATGATCTGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400
Db 2341 ACATGATCTGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400
Qy 2401 GCAGAGAG 2460
Db 2401 GCAGAGAG 2460
Qy 2461 TCGTCTGAACTCTCTTAATCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2520
Db 2461 TCGTCTGAACTCTCTTAATCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2520
Qy 2521 ATCTTTGGGCAATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2580
Db 2521 ATCTTTGGGCAATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2580
Qy 2581 TGTATGTAAG 2640
Db 2581 TGTATGTAAG 2640
Qy 2641 AAAGGGGAACACCT 2700
Db 2641 AAAGGGGAACACCT 2700
Qy 2701 TTGTTAAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2760
Db 2701 TTGTTAAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2760
Qy 2761 ACCCATTTCTGGAACCT 2820
Db 2761 ACCCATTTCTGGAACCT 2820
Qy 2821 TAGGACTTAAGCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2880
Db 2821 TAGGACTTAAGCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2880

RESULT 5
US-08-956-869-11
; Sequence 11, Application US/08956869
; Patent No. 6030808
; GENERAL INFORMATION:
; APPLICANT: Darnell Jr., James E.
; APPLICANT: Schindler, Christian W.
; APPLICANT: Fu, Xian-Yuan
; APPLICANT: Wen, Zilong
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,869

1501 TGTGGTATAACATGCTGACCAATTAACCCCAAGACGTAAGCTTCTTCACTAAGCCGCCAA 1560
1501 TGTGGTATAACATGCTGACCAATTAACCCCAAGACGTAAGCTTCTTCACTAAGCCGCCAA 1560
1561 TTGGAACCTGGAGCAAGTGGCCGAGGTGCTCAAGTGGCAGTTCTGCTCAGCACAAGC 1620
1561 TTGGAACCTGGAGCAAGTGGCCGAGGTGCTCAAGTGGCAGTTCTGCTCAGCACAAGC 1620
1621 GAAGGCTGACATGACGACCTGACCAAGCTGGGTGAGAGTCTGAGGCGCTGATGTA 1680
1621 GAAGGCTGACATGACGACCTGACCAAGCTGGGTGAGAGTCTGAGGCGCTGATGTA 1680
1621 GAAGGCTGACATGACGACCTGACCAAGCTGGGTGAGAGTCTGAGGCGCTGATGTA 1680
1681 ACTACTGAGGCTGACATGACATGAGGCTAAATTTCTGCAAGAAACATGCTGGCAAGG 1740
1681 ACTACTGAGGCTGACATGACATGAGGCTAAATTTCTGCAAGAAACATGCTGGCAAGG 1740
1741 GCTTCTCCTCTGGGTGCTGAGTATGATGATGATGATGATGATGATGATGATGATGATG 1800
1741 GCTTCTCCTCTGGGTGCTGAGTATGATGATGATGATGATGATGATGATGATGATGATG 1800
1801 CCTTTGGAATGAAGGCTACATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
1801 CCTTTGGAATGAAGGCTACATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
1861 TAACCAAAAGCCCGGCGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
1861 TAACCAAAAGCCCGGCGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
1921 GGGTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
1921 GGGTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
1921 GGGTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
1981 AGCCATACACCAACAGCAGTGAACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
1981 AGCCATACACCAACAGCAGTGAACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
2041 AGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100
2041 AGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100
2101 CCAAGGAGGAGGATTTGGAAAGTACTAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGG 2160
2101 CCAAGGAGGAGGATTTGGAAAGTACTAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGG 2160
2161 ACCCAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
2161 ACCCAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
2161 ACCCAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
2221 GCAGCAATACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2280
2221 GCAGCAATACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2280
2221 GCAGCAATACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2280
2281 GAATTAACGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
2281 GAATTAACGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
2281 GAATTAACGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
2341 ACATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
2341 ACATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
2401 GCAGAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2460
2401 GCAGAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2460
2461 TCGTCTGAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2520
2461 TCGTCTGAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2520
2461 TCGTCTGAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2520
2521 ATCTTTGGGCAATCTGGGCACTTTTAAAGAGAGAAATGAGTGTGGGTGATAAC 2580
2521 ATCTTTGGGCAATCTGGGCACTTTTAAAGAGAGAAATGAGTGTGGGTGATAAC 2580
2521 ATCTTTGGGCAATCTGGGCACTTTTAAAGAGAGAAATGAGTGTGGGTGATAAC 2580
2581 TGTATGTAAG 2640

2581 TGTATGTAAG 2640
2641 AAAGGGAACACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700
2641 AAAGGGAACACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700
2701 TTGTTAGACAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
2701 TTGTTAGACAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
2761 ACCCCATTTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2820
2761 ACCCCATTTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2820
2821 TAGGACTTAAGCCGAGAGGTTCTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 2869
2821 TAGGACTTAAGCCGAGAGGTTCTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 2869

RESULT 6
US-08-948-547-11
Sequence 11, Application US/08948547
Patent No. 6124118
GENERAL INFORMATION:
APPLICANT: Darnell Jr., James E.
APPLICANT: Schindler, Christian W.
APPLICANT: Fu, Xian-Yuan
APPLICANT: Wen, Zhong
TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauder & Jackson
STREET: 111 Harkensack Avenue
CITY: " kensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948,547
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/212,185
FILING DATE: 11-MAR-1994
APPLICATION NUMBER: US 07/980,498
FILING DATE: 23-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/854,296
FILING DATE: 19-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO US93/02569
FILING DATE: 19-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,588
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-073 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:

LENGTH: 2869 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: mouse
IMMEDIATE SOURCE:
LIBRARY: splenic/thymic
CLONE: Murine 19sf6
FEATURE:
NAME/KEY: CDS
LOCATION: 69..2378
US-08-948-547-11

Query Match 100.0%; Score 2869; DB 3; Length 2869;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCCGACACCGACGGCCGACAGTGGGCTCAGCCCGGAGACAGTGCAGACCCCTGACT 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 GCCCGACACCGACGGCCGACAGTGGGCTCAGCCCGGAGACAGTGCAGACCCCTGACT 60
OY GCACAGAGATGCTCAGTGGAAACAGCTGCAGAGCTGCAGACACCGTACTGGAAGCAGC 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 GCACAGAGATGCTCAGTGGAAACAGCTGCAGAGCTGCAGACACCGTACTGGAAGCAGC 120
OY 121 TGCACACAGCTGTACAGCGACACGTTCCCATGAGAGCTGCGGACATTCCTGGACCTTGA 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 TGCACACAGCTGTACAGCGACACGTTCCCATGAGAGCTGCGGACATTCCTGGACCTTGA 180
OY 181 TTGAGAGTCAGACTGGGCGATGATGCGCAGCAAGAGTGCATGACCAAGTGTGTTTC 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 TTGAGAGTCAGACTGGGCGATGATGCGCAGCAAGAGTGCATGACCAAGTGTGTTTC 240
OY 241 ATAAATCTCTGGTGAAATTTGACCAAGCAATATAGCCGATTCCTGCAAGAGTCCATATGCC 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 ATAAATCTCTGGTGAAATTTGACCAAGCAATATAGCCGATTCCTGCAAGAGTCCATATGCC 300
OY 301 TCTATCAGACACACCTTGAAGAAATCAAGCAATTTCTGCACAGCAGATATCTTGAAGAGC 360
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 TCTATCAGACACACCTTGAAGAAATCAAGCAATTTCTGCACAGCAGATATCTTGAAGAGC 360
OY 361 CAATGGAATTTGCCCGATCGTGCCCGCATGCTGTGGGAAGAGTCTGCTCCAGCA 420
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 CAATGGAATTTGCCCGATCGTGCCCGCATGCTGTGGGAAGAGTCTGCTCCAGCA 420
OY 421 CGGAGCGCACGGCAGCCAGCAAGGGGGCCAGGCCAACCCCAACAGCCCGCTAGTGA 480
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 421 CGGAGCGCACGGCAGCCAGCAAGGGGGCCAGGCCAACCCCAACAGCCCGCTAGTGA 480
OY 481 CAGAGAGCAGCAATGTTGGAGAGCATCTTCAGAGATGTCCGGAAGGAGATGAGAGATC 540
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 481 CAGAGAGCAGCAATGTTGGAGAGCATCTTCAGAGATGTCCGGAAGGAGATGAGAGATC 540
OY 541 TAGAAGCAAAATGAAGGTGTGGAGAACTTCAGAGAGCACTTTGATTCAACTACAAAA 600
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 541 TAGAAGCAAAATGAAGGTGTGGAGAACTTCAGAGAGCACTTTGATTCAACTACAAAA 600
OY 601 CCCCTCAAGAGCAGCAAGAGACATGCAAGATGTGAAACCAACAGTCTGTGACACAGAC 660
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 601 CCCCTCAAGAGCAGCAAGAGACATGCAAGATGTGAAACCAACAGTCTGTGACACAGAC 660
OY 661 AGAAGATCAGCAGCTGGAACAGATGCTCAGAGCCTTGAGCAGATGGGAGAGCATTTG 720
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 661 AGAAGATCAGCAGCTGGAACAGATGCTCAGAGCCTTGAGCAGATGGGAGAGCATTTG 720
OY 721 TGAGTGAAGTGGCGGGCTTTGTGAGCAATGAGATGCTGAGAGAGCACTGACTGATG 780
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 721 TGAGTGAAGTGGCGGGCTTTGTGAGCAATGAGATGCTGAGAGAGCACTGACTGATG 780

OY 781 AAGAGCTGGTGAAGAGAGGGCCAGAGATGCGGTGATGAGAGCCCTCCACACA 840
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 781 AAGAGCTGGTGAAGAGAGGGCCAGAGATGCGGTGATGAGAGCCCTCCACACA 840
OY 841 TCTGCTGGACCGCTGTGGAAACCTGATTAATCTTATAGCAGAAATCTCAACTCAGACCC 900
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 841 TCTGCTGGACCGCTGTGGAAACCTGATTAATCTTATAGCAGAAATCTCAACTCAGACCC 900
OY 901 GCCAACAATAATTAAGAACTGAGAGAGCTGCAGCAAGAAAGTGTCTACAAAGCGGACCTTA 960
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 901 GCCAACAATAATTAAGAACTGAGAGAGCTGCAGCAAGAAAGTGTCTACAAAGCGGACCTTA 960
OY 961 TCGTGCAGACCGGCGCCATGCTGAGAGAGAGATGCTGAGAGCTTTGAGAACTTAATGA 1020
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 961 TCGTGCAGACCGGCGCCATGCTGAGAGAGAGATGCTGAGAGCTTTGAGAACTTAATGA 1020
OY 1021 AGAGTGCCTTCGTGGTGGAGCGGACCGCTGCATGCCATGCAACCGGACCGGCTTGA 1080
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1021 AGAGTGCCTTCGTGGTGGAGCGGACCGCTGCATGCCATGCAACCGGACCGGCTTGA 1080
OY 1081 TCATCAGAGACTGTGTCCAGTTTACAGAAAGTCAAGTGTGCTGCAATTTCTCGAGT 1140
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1081 TCATCAGAGACTGTGTCCAGTTTACAGAAAGTCAAGTGTGCTGCAATTTCTCGAGT 1140
OY 1141 TGAATTTATCAGCTTAAATTTAAAGTGTGATTAAGAGCTGTGGGATGTTGCTGCC 1200
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1141 TGAATTTATCAGCTTAAATTTAAAGTGTGATTAAGAGCTGTGGGATGTTGCTGCC 1200
OY 1201 TCAGAGGCTCTCGGAATTTTAACATTTCTGGGACAGAACACAAAAGTATATACATGAGAG 1260
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1201 TCAGAGGCTCTCGGAATTTTAACATTTCTGGGACAGAACACAAAAGTATATACATGAGAG 1260
OY 1261 AGTCTAACAGCGGACGCTGCTCAGAGTTCAAGACACCTGAGCCTTAGGAGCAGAGAT 1320
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1261 AGTCTAACAGCGGACGCTGCTCAGAGTTCAAGACACCTGAGCCTTAGGAGCAGAGAT 1320
OY 1321 GTGGGAATGAGAGCGCGTGCACATTTGTATGCTCTTGTATGCTGAGAGCTGCACC 1380
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1321 GTGGGAATGAGAGCGCGTGCACATTTGTATGCTCTTGTATGCTGAGAGCTGCACC 1380
OY 1381 TGATACACTTCAGAGCTGAGAGTGTGACCAAGGCGCTCAAGATGTAGACTGAGAGCCACT 1440
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1381 TGATACACTTCAGAGCTGAGAGTGTGACCAAGGCGCTCAAGATGTAGACTGAGAGCCACT 1440
OY 1441 CCTTGGCAGTTTGTGATGCTCCAAACATCTGTCAAGATGCCAAATGCTTGGGCATCAATCC 1500
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1441 CCTTGGCAGTTTGTGATGCTCCAAACATCTGTCAAGATGCCAAATGCTTGGGCATCAATCC 1500
OY 1501 TGTGTTAATACATGCTGACCAATTAACCCCAAGAACGTGAATCTTCACTAAGCCGCCAA 1560
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1501 TGTGTTAATACATGCTGACCAATTAACCCCAAGAACGTGAATCTTCACTAAGCCGCCAA 1560
OY 1561 TTGGAACCTGGAGACCAAGTGGCCGAGGTGCTCACTGGCAGATTTCTGTCACCAACCAAGC 1620
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1561 TTGGAACCTGGAGACCAAGTGGCCGAGGTGCTCACTGGCAGATTTCTGTCACCAACCAAGC 1620
OY 1621 GAGGCGTGAGCAGTGCAGCAGCTGACCAACGCTGAGAGAGCTCTGAGGCGCTGGGTGA 1680
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1621 GAGGCGTGAGCAGTGCAGCAGCTGACCAACGCTGAGAGAGCTCTGAGGCGCTGGGTGA 1680
OY 1681 ACTACTCAGGCTGCATGATCATGAGGCTAAATCTGCAAGAAAGAAACATGGCTGGCAAG 1740
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1681 ACTACTCAGGCTGCATGATCATGAGGCTAAATCTGCAAGAAAGAAACATGGCTGGCAAG 1740
OY 1741 GCTTCTCTCTCTGGGTCTGGCTAGACAAATATCATGCACTTTGTGAAAAAGTATATCTTGG 1800
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1741 GCTTCTCTCTCTGGGTCTGGCTAGACAAATATCATGCACTTTGTGAAAAAGTATATCTTGG 1800
OY 1801 CCTTTTGAATGAAGGTTACTCATGSGTTTCATCAGCAGAGAGCGGAGCGGCATGTC 1860
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1801 CCTTTTGAATGAAGGTTACTCATGSGTTTCATCAGCAGAGAGCGGAGCGGCATGTC 1860

```

OY 1861 TAAGACAAAACCCCGGACCTTCTACTGCGCTTCAAGCAGACGACGAAAGAG 1920
    |||||||
DB 1861 TAAGACAAAACCCCGGACCTTCTACTGCGCTTCAAGCAGACGACGAAAGAG 1920
OY 1921 GGGTCACTTTACCTTGGTGGAAAGACATCAGTGGCAAGACCCAGATCCACTGTAG 1980
    |||||||
DB 1921 GGGTCACTTTACCTTGGTGGAAAGACATCAGTGGCAAGACCCAGATCCACTGTAG 1980
OY 1981 AGCCATACCAAGACAGACAGCTGACATGTCATTTGCTGAAATCATATGCGCTATA 2040
    |||||||
DB 1981 AGCCATACCAAGACAGACAGCTGACATGTCATTTGCTGAAATCATATGCGCTATA 2040
OY 2041 AGATCATGATGAGGACCAACATCTGCTGCTCACTTGTCTACTGCTACCCGACATTC 2100
    |||||||
DB 2041 AGATCATGATGAGGACCAACATCTGCTGCTCACTTGTCTACTGCTACCCGACATTC 2100
OY 2101 CCAAGAGAGAGGCAATTTGAAAGTACTGTAGCCCGAGACCCAGAGACCCCGAACCCG 2160
    |||||||
DB 2101 CCAAGAGAGAGGCAATTTGAAAGTACTGTAGCCCGAGAGCCAGAGACCCCGAACCCG 2160
OY 2161 ACCGAGTAGTGTCTGCGCCGCTACTGAAAGCAAGTTCACTGTGTGACACCAAGCAGCT 2220
    |||||||
DB 2161 ACCGAGTAGTGTCTGCGCCGCTACTGAAAGCAAGTTCACTGTGTGACACCAAGCAGCT 2220
OY 2221 GCAGCAATACCATTTGACCTGCGGATGTCCCGCCGACCTTTAGATTGATGATGAGTTTG 2280
    |||||||
DB 2221 GCAGCAATACCATTTGACCTGCGGATGTCCCGCCGACCTTTAGATTGATGATGAGTTTG 2280
OY 2281 GAAATTAACGCTGAAGGCTGTAGCCCTCAGAGAGAGGCAAGTTGAGTGTGCTGACGTTTG 2340
    |||||||
DB 2281 GAAATTAACGCTGAAGGCTGTAGCCCTCAGAGAGAGGCAAGTTGAGTGTGCTGACGTTTG 2340
OY 2341 ACATGATCTGACCTGCGAGTGTGCTACCTGACCTCCCATGTGAGAGAGTGAACAGAAAGCT 2400
    |||||||
DB 2341 ACATGATCTGACCTGCGAGTGTGCTACCTGACCTCCCATGTGAGAGAGTGAACAGAAAGCT 2400
OY 2401 GCAGAGAGCTGATTTGAGACACCTGCGGCTGCTCACCCCTTAAGCAGCCCAACCCCTA 2460
    |||||||
DB 2401 GCAGAGAGCTGATTTGAGACACCTGCGGCTGCTCACCCCTTAAGCAGCCCAACCCCTA 2460
OY 2461 TCGCTGAAACCTGTAATTTGTGTTGCTCAGATTTTCTTAAATTTCTACTTCTGCT 2520
    |||||||
DB 2461 TCGCTGAAACCTGTAATTTGTGTTGCTCAGATTTTCTTAAATTTCTACTTCTGCT 2520
OY 2521 ATCTTTGGCAATCTGGGCACTTTTAAAGAGAGAAATGAGTGTGGGTGATAAC 2580
    |||||||
DB 2521 ATCTTTGGGCAATCTGGGCACTTTTAAAGAGAGAAATGAGTGTGGGTGATAAC 2580
OY 2581 TGTATTGTAAGAGAGAGACCTCTGAGTCTGGGATGCGGCTGAGACAGAGAGGAGGC 2640
    |||||||
DB 2581 TGTATTGTAAGAGAGAGACCTCTGAGTCTGGGATGCGGCTGAGACAGAGAGGAGGC 2640
OY 2641 AAAGGGGAACACCTCTGCTGCTGCGGCTGCGCTCTCTTTTTCAGACAGCTCGGGGGTTGG 2700
    |||||||
DB 2641 AAAGGGGAACACCTCTGCTGCTGCGGCTGCGCTCTCTTTTTCAGACAGCTCGGGGGTTGG 2700
OY 2701 TTGTTAGACAACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
    |||||||
DB 2701 TTGTTAGACAACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
OY 2761 ACCCATCTGGGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
    |||||||
DB 2761 ACCCATCTGGGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
OY 2821 TAGGACATAAGCCAGAGAGTCTCTTTTAAATTTAAAAAAGAGAGAGAGAGAGAGAGAG 2869
    |||||||
DB 2821 TAGGACATAAGCCAGAGAGTCTCTTTTAAATTTAAAAAAGAGAGAGAGAGAGAGAGAG 2869

```

```

? APPLICANT: James E. Darnell, Jr.
? APPLICANT: Zilong Wen
? APPLICANT: Curt M. Horvath
? APPLICANT: Zhong Zhong
? TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
? NUMBER OF SEQUENCES: 39
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Klauber & Jackson
? STREET: 411 Hackensack Avenue
? CITY: Hackensack
? STATE: New Jersey
? COUNTRY: USA
? ZIP: 07601
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: Patent in Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: PCT/US95/17025
? FILING DATE: 28-DEC-1995
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/369,796
? FILING DATE: 06-JAN-1995
? CLASSIFICATION:
? ATTORNEY/AGENT INFORMATION:
? NAME: Jackson Esq., David A.
? REGISTRATION NUMBER: 26,742
? REFERENCE/DOCKET NUMBER: 600-1-116
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 201 487-5800
? TELEFAX: 201 343-1684
? TELEX: 133521
? INFORMATION FOR SEQ ID NO: 11:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2869 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: both
? TOPOLOGY: unknown
? MOLECULE TYPE: cDNA
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? ORIGINAL SOURCE:
? ORGANISM: Mouse
? IMMEDIATE SOURCE:
? LIBRARY: splenic/lymnic
? CLONE: Murine 19s16
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 69..2378
? PCT-US95-17025-11

Query Match 100.0%; Score 2869; DB 5; Length 2869;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCCCGACACCGACGCGCGGCGGCTGCGCTGAGCCCGGAGACAGTGGAGACCCCTGACT 60
DB 1 GCCCGACACCGACGCGCGGCGGCGGCTGCGCTGAGCCCGGAGACAGTGGAGACCCCTGACT 60
OY 61 GCAGCAGATGGCTCAGTGAACAGCTGCGAGCTGGACACAGCTACTGTAAGCAGC 120
DB 61 GCAGCAGATGGCTCAGTGAACAGCTGCGAGCTGGACACAGCTACTGTAAGCAGC 120
OY 121 TGCACACAGCTGTACAGCAGACGTTCCCATGAGAGCTGGCAGTTCCTGGACCTTGA 180
DB 121 TGCACACAGCTGTACAGCAGACGTTCCCATGAGAGCTGGCAGTTCCTGGACCTTGA 180
OY 181 TTGACAGTCAAGACTGGGCAATATGAGCCAGCAAAAGATCAATCAACGTTGGTGTTC 240
DB 181 TTGACAGTCAAGACTGGGCAATATGAGCCAGCAAAAGATCAATCAACGTTGGTGTTC 240

```

Db 181 TTGAGAGTCAAGACTGGGATATGACAGCAAAAGATGCATGCCACGTTGGTTC 240
OY 241 ATATATCTTGGCTGAATTTGACACCAATATATGCGCATTCGCAAGAGATCCCATGTC 300
Db 241 ATATATCTTGGCTGAATTTGACACCAATATATGCGCATTCGCAAGAGATCCCATGTC 300
OY 301 TCTATCAGCAACACTTCGAAGAATCAAGCAGTTTCTGCAGAGCAGGTATCTTGAGAAC 360
Db 301 TCTATCAGCAACACTTCGAAGAATCAAGCAGTTTCTGCAGAGCAGGTATCTTGAGAAC 360
OY 361 CATGGAATTCGCGGATGCTGGCGGATGCTGGGAGAGTCTGGCTCTCCAGCA 420
Db 361 CATGGAATTCGCGGATGCTGGCGGATGCTGGGAGAGTCTGGCTCTCCAGCA 420
OY 421 CGGACGCCAGCGGACGCCAGCAAGGGGGCGGAGGCCAACGCCAACAGCGCGGTACTGA 480
Db 421 CGGACGCCAGCGGACGCCAGCAAGGGGGCGGAGGCCAACGCCAACAGCGCGGTACTGA 480
OY 481 CAGAGAAAGCAGATGTTGGAGCAGCATCTTGAAGATGTCGGGAACCGGATGAGATC 540
Db 481 CAGAGAAAGCAGATGTTGGAGCAGCATCTTGAAGATGTCGGGAACCGGATGAGATC 540
OY 541 TAGAACGAAATGAAGTGTGTGAGAACCTCCAGAGCAGACTTGTGATTTCACTACAAA 600
Db 541 TAGAACGAAATGAAGTGTGTGAGAACCTCCAGAGCAGACTTGTGATTTCACTACAAA 600
OY 601 CCTCAAGAGCCAGAGACATGACAGATCTGAATGGAACCAACCACTCTGTGACCCAGAC 660
Db 601 CCTCAAGAGCCAGAGACATGACAGATCTGAATGGAACCAACCACTCTGTGACCCAGAC 660
OY 661 AGAAGATGACAGCTGGAACAGATGCTCACACCCCTGACAGCAGATGCGGAGAGCTTG 720
Db 661 AGAAGATGACAGCTGGAACAGATGCTCACACCCCTGACAGCAGATGCGGAGAGCTTG 720
OY 721 TGAATGAGCTGGCGGGGCTTCTTCAAGCAATGAGTACGTGCAAGAACACTGATGATG 780
Db 721 TGAATGAGCTGGCGGGGCTTCTTCAAGCAATGAGTACGTGCAAGAACACTGATGATG 780
OY 781 AAGAGCTGGCTGCTGGAAGAGGCGGCGAGATGCGGTGCTCATGCGGAGCGCTCCCAACA 840
Db 781 AAGAGCTGGCTGCTGGAAGAGGCGGCGAGATGCGGTGCTCATGCGGAGCGCTCCCAACA 840
OY 841 TCTGCTGACCGCTGGAACACTGATTAATTGAGCAGATCTCAACTTCAGACCC 900
Db 841 TCTGCTGACCGCTGGAACACTGATTAATTGAGCAGATCTCAACTTCAGACCC 900
OY 901 GCCAACAATTAAGAACTGAGAGCTGCGAGCAAGAAAGTGTCTCAAAAGGAGCCCTTA 960
Db 901 GCCAACAATTAAGAACTGAGAGCTGCGAGCAAGAAAGTGTCTCAAAAGGAGCCCTTA 960
OY 961 TCGTGCAGCAGCGGCCCATGCTGAGAGAGAGATGCTGAGAGCTGTTCAAGAACTTAATGA 1020
Db 961 TCGTGCAGCAGCGGCCCATGCTGAGAGAGAGATGCTGAGAGCTGTTCAAGAACTTAATGA 1020
OY 1021 AAGAGTCTTCTGTGTGTGAGAGCGGCGGAGCTGCAATGCCATGCAACCGGAGCCCTTAG 1080
Db 1021 AAGAGTCTTCTGTGTGTGAGAGCGGCGGAGCTGCAATGCCATGCAACCGGAGCCCTTAG 1080
OY 1081 TCAATCAAGAGCTGCTGCACTTACCAAGAAATGCAAGTGTGCTGCTCAAAATTTCTGAGT 1140
Db 1081 TCAATCAAGAGCTGCTGCACTTACCAAGAAATGCAAGTGTGCTGCTCAAAATTTCTGAGT 1140
OY 1141 TGAATATACAGCTTAAATTAAGTGTGATTAAGATCTGTGGGATGTTGCTGCC 1200
Db 1141 TGAATATACAGCTTAAATTAAGTGTGATTAAGATCTGTGGGATGTTGCTGCC 1200
OY 1201 TCGAGAGGCTCTGGAATTTAATCTGGGACGAACAAAGATGATGAATGAGG 1260
Db 1201 TCGAGAGGCTCTGGAATTTAATCTGGGACGAACAAAGATGATGAATGAGG 1260
OY 1261 AGTCTAACACAGGCGAGCTGTCTGCAAGATTTCAAGCACTGACCTTTAGAGCAGAGAT 1320
Db 1261 AGTCTAACACAGGCGAGCTGTCTGCAAGATTTCAAGCACTGACCTTTAGAGCAGAGAT 1320

OY 1321 GTGGAAATGAGGCGGCTGCCAATTTGATGCTCTTGTATGCTGACTGAGAGCTGCACC 1380
Db 1321 GTGGAAATGAGGCGGCTGCCAATTTGATGCTCTTGTATGCTGACTGAGAGCTGCACC 1380
OY 1381 TGATCAGCTTGAAGACTGAGAGTGTACCAAGAGCCCAAGATTGACTTGAAGACCCACT 1440
Db 1381 TGATCAGCTTGAAGACTGAGAGTGTACCAAGAGCCCAAGATTGACTTGAAGACCCACT 1440
OY 1441 CTTGGCAGTTGGTGTATCTCCAAACTGTGTAGATGCCAAATGCTTGGGCAATATCC 1500
Db 1441 CTTGGCAGTTGGTGTATCTCCAAACTGTGTAGATGCCAAATGCTTGGGCAATATCC 1500
OY 1501 TGTGTATTAACATGCTGACCAATTAACCCAGAAAGTGAACCTTCTCACTAAGCCGCAA 1560
Db 1501 TGTGTATTAACATGCTGACCAATTAACCCAGAAAGTGAACCTTCTCACTAAGCCGCAA 1560
OY 1561 TTGGAACCTGGAGCCAAAGTGGCGGAGGTGCTCAGCTGGCAGTTCTGTCCAGCCAAAGC 1620
Db 1561 TTGGAACCTGGAGCCAAAGTGGCGGAGGTGCTCAGCTGGCAGTTCTGTCCAGCCAAAGC 1620
OY 1621 GAGGCTGAGCATGAGACAGCTGACAAAGCTGTGAGAACTCTTAGGCTGTGTGA 1680
Db 1621 GAGGCTGAGCATGAGACAGCTGACAAAGCTGTGAGAACTCTTAGGCTGTGTGA 1680
OY 1681 ACTACTAGAGGTGTACAGATACATGGGCTAAATTTCTGCAAGAAACATGCTGGCAGG 1740
Db 1681 ACTACTAGAGGTGTACAGATACATGGGCTAAATTTCTGCAAGAAACATGCTGGCAGG 1740
OY 1741 GCTTCTCTCTGTGGTGTGCTGACAAATATGCACTGCACTTGTGAAGAAATATATCTTG 1800
Db 1741 GCTTCTCTCTGTGGTGTGCTGACAAATATGCACTGCACTTGTGAAGAAATATATCTTG 1800
OY 1801 CCTTTTGAATGAAGGCTATCATGCTGTTTCAATCAGCAAGAGCGGAGCCGACATCC 1860
Db 1801 CCTTTTGAATGAAGGCTATCATGCTGTTTCAATCAGCAAGAGCGGAGCCGACATCC 1860
OY 1861 TTAAGCAAAAGCCCGGAGACCTTCTACTGCGCTTACAGCAGAGCAGCAAGAAAGAG 1920
Db 1861 TTAAGCAAAAGCCCGGAGACCTTCTACTGCGCTTACAGCAGAGCAGCAAGAAAGAG 1920
OY 1921 GGTCTACTTCTACTTGGTGGTGAAGAGCATGATGAGGAGACCCAGATCTGTAG 1980
Db 1921 GGTCTACTTCTACTTGGTGGTGAAGAGCATGATGAGGAGACCCAGATCTGTAG 1980
OY 1981 AGCCATACAGCAAGCAGCAGCTGACCAACATGCTATTTGCTGAATCATATGCGCTATA 2040
Db 1981 AGCCATACAGCAAGCAGCAGCTGACCAACATGCTATTTGCTGAATCATATGCGCTATA 2040
OY 2041 AGATCATGATGAGCAGCAACATGCTGTCTCAGCTTGTCTACCTTACCCGACATTC 2100
Db 2041 AGATCATGATGAGCAGCAACATGCTGTCTCAGCTTGTCTACCTTACCCGACATTC 2100
OY 2101 CCAAGAGAGAGCATTTTGAAGTACTGTAGGCCGAGAGCCAGAGACCCCGAAGCCG 2160
Db 2101 CCAAGAGAGAGCATTTTGAAGTACTGTAGGCCGAGAGCCAGAGACCCCGAAGCCG 2160
OY 2161 ACCCAGTAGTGTGCGCCCTTACCTGAGAGCAAGTTCATCTGTGTACACCAAGCCT 2220
Db 2161 ACCCAGTAGTGTGCGCCCTTACCTGAGAGCAAGTTCATCTGTGTACACCAAGCCT 2220
OY 2221 GCAAGCAATACATTTGACCTGCGATGCTCCCGCAGCTTTAGATTCATTTGATGACGTTTG 2280
Db 2221 GCAAGCAATACATTTGACCTGCGATGCTCCCGCAGCTTTAGATTCATTTGATGACGTTTG 2280
OY 2281 GAAATTAAGGTGAAGGTGCTGAGCCCTCAGCAGAGAGGCGATTGATGCTGCTACGTTTG 2340
Db 2281 GAAATTAAGGTGAAGGTGCTGAGCCCTCAGCAGAGAGGCGATTGATGCTGCTACGTTTG 2340
OY 2341 ACATGATCTGAGCTCGGAGTGTGCTACCTGCCCAATGTAGAGAGCTGAACCAAGCT 2400
Db 2341 ACATGATCTGAGCTCGGAGTGTGCTACCTGCCCAATGTAGAGAGCTGAACCAAGCT 2400

[illegible]

```

RESULT      8
; Sequence 82, Application US/09288461
; Patent No. 6159694
; GENERAL INFORMATION:
; APPLICANT: Karrias, James G.
; TITLE OF INVENTION: Antisense Oligonucleotide Modulation of STAT3
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0338
; CURRENT APPLICATION NUMBER: US/09/288,461
; CURRENT FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 82
; LENGTH: 2869
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (69)..(2381)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: U06922
; DATABASE ENTRY DATE: 1994-07-01
US-09-288-461-82

```

Query Match	Similarity	99.8%	Score 2864.2	DB 3	Length 2869
Best Local	Similarity	99.9%	Pred. No. 0		
Matches 2866	Conservative	0	Mismatches 3	Indels 0	Gaps 0
QY	1	GGCGGACGACCGCCGCGCAGTGGGCTCAGCCCGAGACAGTCGAGACCCCTGACT	60		
Db	1	ggcgcgaccagcagcgccggccagltgggtctagcccggaagacagtcgagacccttgact	60		
QY	61	GGAGCGAGATGGGCTCGTGGAGCACACTCTGCACAGCTGGACACAGCTCTCCGTAAGCAGC	120		
Db	61	ggagcgaggttggtctcagttggaaccagctcgcagcagcttgacaacacgctcccggaagcagc	120		
QY	121	TGCACGAGCTGTACAGCAGACAGCTTCCCATGGAGCTGGGGCAGTTCTCGGACACCTTGA	180		

Db	121	tgacccagcgtgacagcgacacglltccca tggagctcgcggaagcttcttgcaccttgg	180
Oy	181	TTGAGAGTCAAGACCTGGCATATGACGCGCAGCAAGAGTGCATGCGCCAGTTGGTGTTC	240
Db	181	tggagatcaagacttggcatatgcagccagaagaagctcatatgcacgttgglltllc	240
Oy	241	ATATATCTCTTGGGGAATTTGACACCATATATAGCGGATTCCTCGCAAGTCCAAATCTCC	300
Db	241	ataatctcttgggtgaatattgaccagcaatataagccgatctcttgaagatgccaatgtcc	300
Oy	301	TCATATCAGCAACACTTTCGAAATATCAGCAAGTTCCTGCAGACAGGTATCTTCAGAAC	360
Db	301	tctatcagcaacaaccttcgaagaatcaagcaatllcttgcgaagcagttalcttgaagac	360
Oy	361	CAATGCAAAATTTGGCCGGATCTGTGCGCCGATGGCTGTGGGAACATGCTCGCCCTCCAGA	420
Db	361	caatggaatatgtcccgga tccgtggccgagatgccgttgggaagaagcttcgtcctccaga	420
Oy	421	CGGACGCCACGCGCAGCCAGCAAGGGGGCCAGGCCAACCCACCAACAGCTCGCTAGTGA	480
Db	421	cggaagccacggaagcccgagccaaggggcccagggccaaaccaaccagcgcgtatgta	480
Oy	481	CAGAGAACGACGACAGATGTTTGAGACAGCATCTTCAGAGATGCCGGAAGGAGATGCAGATC	540
Db	481	cagagaagcagcagatgtcttggagcaacatcttcaagatgtccggaagcgatgtcgaagatc	540
Oy	541	TAGAAACGAAATATGAAGCTGTGTGACAACTCCAGACACCTCTTATTTCAACTACAAA	600
Db	541	tagaaacgaaaaatgaaggtgtgtggagaaacctccagagcagctttagtttcaactacaaa	600
Oy	601	CCCTCAAGAGCCAAAGAGACATGCAGAGATCTGAATGGAACAACCAATCTGTGACACAG	660
Db	601	ccctcaagagccaaagagagacatgcagatctgaattgaaacaaccaagcttgtgacagac	660
Oy	661	AGAGATGSCACCACTTCGAGACACATGCTCACACCTCGGACACAGATGGCGGAAGACATTG	720
Db	661	agaagatgtcaacagacttggaaacagatgtctcaacgcccgtgaaacagatgtcgaagaagcatlgt	720
Oy	721	TGAGTGACCTGGCGCGGCTCTTCTTGACCAATGGAGTACGTGACAGAACTGACTGATG	780
Db	721	tgaigtgacttggcgcggtcttcttgcagcaattggagatgtgacagaaactgactgatg	780
Oy	781	AAGACTTGCGTGAATGGGAAGAGCGGCTGCAGATGCGCGTGCATTCGAGGCGCTCCCAAA	840
Db	781	aagacttgcgtgaatgggaagagcggtgcagatgtgcgtgtcatgtggagggccctcccaaa	840
Oy	841	TCTGCTGTGACCGCTGTGGAAACCTGTGATACTTCATTACCAAACTCAACTTCAGACCC	900
Db	841	ctctgcttggacgcgtcttggaaacctgtgaatacttcatltagcagaatctcaactcaagccc	900
Oy	901	GCCAAACAATTTAAGAACTGAGAGAGCTGCAGCAAACTGTCTTACAAGGGCGACCTTA	960
Db	901	gccaaacaatttaagaaacttggagagctgtcagaagaagaatgtcttcaagaaggctgacctta	960
Oy	961	TGCTGCACACACCGGCCCATGTGTGGAGAGAGATCTGTGAGGCTTCACAGAACTTAATGA	1020
Db	961	tctgtgcacacacggcccatgtctggagagagagatctgttgaagctgttccgaacttaatga	1020
Oy	1021	AGAGTGCCTTGTGTGTGAGCGGACGCGCTGCATGTCGCCATGCACCCGAGCGGCCCTTAG	1080
Db	1021	agagtgcttctgtgtgtgtgagcgagccctgtcatgtccatgcacccgagccggcccttag	1080
Oy	1081	TCATCAACACGTGTGTGAGTTTACCGCAAGAACTCGATGTGGTGCGTCAAAATTCCTGTAGT	1140
Db	1081	tcatcaaacacgtgtgtgagtttaccgcaagaaactcgaatgttgcgttgcgaatctccaggt	1140
Oy	1141	TGAATTTACACTTAAAAATTTAAAGTGTGCAATTTGATTAAGAACTCTGGGCAATGTTGCTGCC	1200
Db	1141	tgaatttatcactttaaatttaaaatgtgtgcatatgaagaactctgggagttgtgtgcctc	1200
Oy	1201	TCAGA TGGTCTTCGCAAAATTTAACATTTGTGGGCGCAGCAACCAAAAAGTATGAAATCATGAGAG	1260
Db	1201	tcaagtgcttcttcggaatattcaacttcttggcgcagacaacaaagtgatgaacatvgagag	1260

QY	1261	AGTTAAACAAGCAGAGCCTTCTGCGAGTTCAAGACTGACCTTGGGACGAGAT	1320
Db	1261	agttcaacaagagagcctcttcgcagagttccaagaccctcttgaggagcaagat	1320
QY	1321	GTGGGAATGAGAGCGCGTCCCAATTGTGATGCTCCTTGATCTGACTGAGAGAGCTGCACC	1380
Db	1321	gtgggaatgagagcgctgcacaatttgatgagctctcttgactgagtagagagctgcacc	1380
QY	1381	TGATACACCTTGAGAGACTGAGCTGTACCAACGAGCGCTCAAGTTGACCTAGAGACCCACT	1440
Db	1381	tgatcaaccttgagagctgaggtgtaccaccaagagctcaagttgacctagagaccact	1440
QY	1441	CGTTGCCAGTTGTGTGATCTCCACACATCTGTACATGTCACCAATCCTTGGCATCATATCC	1500
Db	1441	cgttgccagttgtgtgattctccaacatctgtcagaatgccaatgcttggcatalcatcc	1500
QY	1501	TGTGGTAAACATCTCTACCAATTAACCCCAAGACAGCTGAATCTTCACTAAGCCGCCAA	1560
Db	1501	tgtgtataacatctctgacccaataaccccaagaagcttgaccttctcaataagccgcmaa	1560
QY	1561	TTTGAACTCGGGAGACACAGTGGCCGAGGTGCTACAGCTGGCAGTTCCTGTCTCACTACCAAC	1620
Db	1561	tttgaaactcgggagacagtgccgaggtgctcagctgagcttctgtccaccacaagc	1620
QY	1621	GAGGCGCTGAGCATGCAGCACCTGACAAACGCTGAGCTAGAACGCTCTAGAGGCTGTGTGA	1680
Db	1621	gaggcgctgagcatgcagcacctgacaaacgctgagctagaaagctcttaggacctgtgtga	1680
QY	1681	ACTACTCAGGCTGTGACATACATGGGCTTAATTTGCAACAACAAACATGGCTGGCAAG	1740
Db	1681	actactcaggctgtgacatacatgggcttaatttgcaacaacaaacatggctggcaag	1740
QY	1741	GCTTCTCCTTCTGGGCTGGCTAGAACAAATCATCATCACCCTTGTGAAAAAGTATATCTTGG	1800
Db	1741	gcttctccttctgggctggctagaaacaaatcatcatcacccttgtgaaaaagtatatcttg	1800
QY	1801	CCCTTTGGAATGAAGGTTACATCATGGGTTTATCAGCAGAGAGCGGAGCGGCCATTC	1860
Db	1801	ccctttggaatgaaggttacatcatgggtttatcagcagagagcgggagcggccattcc	1860
QY	1861	TAAACACAAACCCCGCGGACCTTCTCTACTCGGCTTGACGCAGAGCAGCAACGAAGAG	1920
Db	1861	taaacacaaaccccgcgacaccttctactctgcttcaggaagcagcaagaagaag	1920
QY	1921	GGGTGACTTTCATCTGGGTGGGAAAGACATCAGTGGCAAGCCGAGATCCAGTTGTAG	1980
Db	1921	gggtgacttctcatcttgaggaaagacatcagtggaagaccagatccagttgtag	1980
QY	1981	AGGCATACACCAAGCACAGCTAAGACAGATGTTCATTTGCTTAATATCATATGGCTATA	2040
Db	1981	aggcatacaccaagcacagctaaagacagtgttcataatgcttaataatcatatggctata	2040
QY	2041	AGATCATGAGTGGCAGCAACATCTGTGTCTTCACATTTCTACTCTACCCGACATTC	2100
Db	2041	agatcatagtagcagccaacatctgtgtcttcaatcttcaactctlaacccgcgacatc	2100
QY	2101	CCAAGGAGAGGCAATTTGAAAGTACTGTAGAGCCGAGACCAGAGACACCCGCAAGCCGC	2160
Db	2101	ccaaggagagagcatcttgaaagtactltagcgcgagagcagacacccgcgaagcg	2160
QY	2161	ACCCAGGATGTCGTCGCCCTACTCTAAGAACCAAGTTCATCTGTGTGACACCAAGACCT	2220
Db	2161	acccagatgtagtcgctgccttacctgaagacaagttcatctgttgacaccaagacct	2220
QY	2221	GCAGCAATACCATTTGACCTCCCATGTCTCCCGCCGCACTTTAGATTCATTTGATGACGTTG	2280
Db	2221	gcagcaatatccattgacctcgcgatgtcccccgcaccttagatctcatgtatgacglttg	2280
QY	2281	GAAATTAACGGTGAAGGTGCTGACCTCTACAGAGAGGAGCAGTTTATGTCGCTCAGCTTGG	2340
Db	2281	gaataatacggtagaggtgctgagcctctcagcagagagagttttagtgcgtccaaaglttg	2340

OY	2341	ACATGAGTCTGACCTCGGAGGTGGTACTCCCATGTCAGAGACTGAAACCAACT	2400
Dd	2341	acatgagctcgcgcccgaggtgctactcccccagtggaggctgaaccataagct	2400
OY	2401	GCAGAGAGTGACTTTGAGACACACTCCCCGTGCTCCACCCTTAACAGCCGAACCCATA	2460
Dd	2401	gcagagagcttjacttgagaaaccttcgcccgctcacccccctaagcagcgaaacccata	2460
OY	2461	TGCTGTGAACCTCCTAACTTTGTGGTCCAGATTTTTTTTTAAATTCCTACTTCTACT	2520
Dd	2461	tgccttgaaacctccaactacttggctccagaattttttttaattctcacttcyct	2520
OY	2521	ATCTTGGGCAATCTGGGCACCTTTTAAAGAAGAAATAGTAGTGAGTGAGTGAATAAC	2580
Dd	2521	atcttgggcaatcctgggcaccttttaaaaaggagaalagtagtgtgtggtaaac	2580
OY	2581	TGTTATGTAAGAGAGAGAGACTCTGAGTCTGGGATGGGCTGAGAGACAGAGGAGGC	2640
Dd	2581	tgltatgtaaaagagagaaacctctgagctcgtggagtggggtctgagagaaggaagc	2640
OY	2641	AAAGGGGAAACACTCTCTGCTGCCGCCCTGCGCTTCCTTTTACACAGCTCGGGGTTGG	2700
Dd	2641	aaaggggaacaacctctctgtcccgccctgcgcccttttctagcagctcgggggtlrg	2700
OY	2701	TTGTTAGACAAGTGCCTTCTCTGTGCCCATGCTACCTGTGGCCCACTCTGTGACTGAT	2760
Dd	2701	ttgttagucaagtgctcctcgtgcceatgctactctgtgcaccactctgagctgal	2760
OY	2761	ACCCGATTCGTGGAGACTCCTGGCTCGACCTTCMAACCTGTGTAAATPNCATPNCATGAAGC	2820
Dd	2761	accccatcttggaacctccctgcgcctccacttcaacctctgctaatalatcacatagaagc	2820
OY	2821	TAGGACTAAGCCCAGAGAGGTTCTCTTTAAATTAATAAAAAAAAAAAAAA	2869
Dd	2821	taggacttaagccaggaaggtctctctttaaatlaaaaaaaaaaaaaa	2869
RESULT 9			
US-09-364-970-9			
; Sequence 9, Application US/09364970			
; Patent No. 6235873			
; GENERAL INFORMATION:			
; APPLICANT: Bromberg, Jacqueline			
; TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR			
; TITLE OF INVENTION: USES FOR IDENTIFYING MODULATORS OF ACTIVITY INCLUDING			
; FILE REFERENCE: 600-1-252			
; CURRENT APPLICATION NUMBER: US/09/364,970			
; CURRENT FILING DATE: 1999-07-31			
; NUMBER OF SEQ ID NOS: 10			
; SOFTWARE: Patentin Ver. 2.0			
; SEQ ID NO 9			
; LENGTH: 2869			
; TYPE: DNA			
; ORGANISM: Mus musculus			
US-09-364-970-9			
Query Match 99.7%; Score 2861; DB 4; Length 2869;			
Best Local Similarity 99.8%; Pred. No. 0;			
Matches 2864; Conservativity 0; Mismatches 5; Indels 0; Gaps 0;			
OY	1	GCGGAGCACCGACGCGCGCGCCAGTCCGGCTTAGCCCCGAGACACTGCAGAACCCCTGACT	60
Dd	1	gccgagaccagccagcgccgagctcggctcagcccggaagacagtcgagaccctgact	60
OY	61	GCAGCAGAGTAGCTCTAGTGGAAACGAGCTGCACACACTGGACACAGAGCTAACGTAAGCAGC	120
Dd	61	gcagcagagtagctcagtggaacacgctcgcagcagctggaacacagctclactctgaagcagc	120
OY	121	TGCACCACTGTACACGCGACACAGCTTCCCAATGAGACTGGCGAGCATTTCTGGCACTTTGA	180
Dd	121	tgcaccaactgtlaacgagacagcttcccaatgagctgtagcgaagcttcccttgacacttga	180

181 TTGAGAGTCAAGACTGGGATATGCGAGCAAGAGTGCATGCGACGTTGTTTC 240
181 ttgagagtaagaactgggcatatgcgagcaagaagatcacaatgcgttggtttc 240
241 ATAAATCTTGGTGAAATGAGCAGCAATATAGCCGATTTCGCAAGAGTCCAAATGTCC 300
241 ataattcttgggtgaattgaccagcaatataagccgattctcgcaagaagttccaatgtcc 300
301 TCTATCAGCACAACCTTCCAGAAATATCAGACAGTCTTTCGAGACAGAGTATCTTGACAGC 360
301 tctatcagcaaaccttccgaagaatataagcagttctgcagagcaaglatcttgaagaagc 360
361 CATGGAATTTGGCCGATGCTGGCCGATGCTGTGGAGAGTCTCGCTCTCCAGA 420
361 caatggaatatggccggtatgtggccgatgctgtggagagagcttcgctctccaga 420
421 CGGACGCCAGCAGCCCGACAGAGGGGCGCAGCCCAACCCAGCCGCGGTAGTGA 480
421 cggacgccagcagcccgacagggggcgagcccaaccacacagccgctagtga 480
481 CAGAGAGCAGCAGATGTTGAGCAGCATCTTACGATGTCGGAAGCAGTGCAGAGATC 540
481 cagagagcagcagatgttggagcagcatcttcaggaatgctcggaaagcgagtgagagatc 540
541 TAGAAGCAGAAATGAGCTGTGGAGAACCTCCAGCAGACTTTGATTTCACACTCAAAA 600
541 tagaagcagaaatgaagtggtggagaaacctcaggaagcacttgcattcaactacaaaa 600
601 CCCCAAGAGCAGCAAGGAGATGCGAGATGTAATGGAACCAACACTCTGTGACAGAC 660
601 ccccaagagcagcaagagaaatcgagaaatctgaaatgaatacaacaaatctgtgcagac 660
661 AGAAGATGACAGCAGCTGGAACAGATGCTCACAGCCCTGGACAGATGCGGAGAACATTTG 720
661 agaagatgacagcagctggaaacagatgctcacagccctggacagatgcggaagacattg 720
721 TGAATGAGCTGGCGGGCTCTTGCAGCAATGAGTACGTGTCAGAGACACTGACTGATG 780
721 tgaatgagctggcgggctcttgcagcaatgagtacgtgacagaaagacactgactgattg 780
781 AAGAGCTGGCTGACTGGAGAGGCGGCCAGAGATGCGGTGCATGCGAGGCCCTCCCAACA 840
781 aagagctggctgactggagagggcgccagagatgcggtgcattgcgagggccctcccaaca 840
841 TCTGCTGAGACCGCTGTGGAACACTGGAATCTTATAGCAAGATTCGAATTCAGACACC 900
841 tctgctgagaccgctgtggaaactggataacttattagcagaatctcaacttcagacc 900
901 GCCAACAATTAAGAAATGAGAGAGCTGCAGCAAAAGTGTCTACAGGCGCACCTTA 960
901 gccacaatatlaagaaactggagagagctgcagcagaagaatgtcttaacaaggcgaccta 960
961 TCGTGACAGACCGGCCATGCTGAGAGAGAGATGCTGAGACTGTTCAGAACTTAATGA 1020
961 tcgtgacagaccggcccatgctgagagagagatgctgagactgttcaagaacttaatga 1020
1021 AGAGTGCTTGTGTGTGAGGCGGACCCCTGATGCCCATGCACCGAGCGGCCCTTAG 1080
1021 agagtgcttgtgtgtgagggcgacccctgatgcccatgcacccgagcgcccttag 1080
1081 TCATCAGACTGTGTGTCAGTTTACACGAGAAAGTCAAGTTGCTGCTCAATTTCTGAGT 1140
1081 tcatcagactgtgtgtcagtttacacgagaaagtcaagttgctgctcaatcttcagtgat 1140
1141 TGAATTTACAGCTTAAATTAAGTGTGATTAAGACTGTGCGGATGTTGTGCGCC 1200
1141 tgaattttacagctttaaattaaagtgtgatataagactgtgggagatgtgtgctgcc 1200
1201 TCAGAGGCTCTCGAAATTTAACATTTGCGGCACGACACAAAAGTGTGACATGAGAG 1260
1201 tcagaggtctctcgaaatttaacatcttcggtgcagcaacaaaagtgtgtaacatggag 1260

1261 AGTCTAACACGGCAGCCTGTCTGACAGAGTTCAAGCAGCTTACGACCTTACGAGCAGAGAT 1320
1261 agtcttaacacggcagcctgtctgacagagttcaagcagcttaccgacctttagggacagagat 1320
1321 GTGGGAATGAGCGCGTGGCCAAATTTGTATGCCCTCTTGTGATCGTGAATGAGAGCTTCACC 1380
1321 gtgggaatgagcgcggtggccaaatttgtatgccctcttgtgatcgatgagagcttcgacc 1380
1381 TGAATCAGCTTGCAGACTGAGCTTACCACCAAGGCCCTCAAGATTGACCTTAGAGACCCACT 1440
1381 tgaatcagcttgcagactgagcttaccaccaaggccctcaagattgaccttagagaccact 1440
1441 CCTTTT CAGTTTGTGTGATTTCCAAACATCTGTACAGATTCGCAATTCGTCGCAATTCAC 1500
1441 cctttt cagtttgtgtgatttccaaacatctgtacagatctgcgaatctgcgcaatctac 1500
1501 TGTGATTAACATGCTGACCAAAATTAACCCCAAGAACGTGAACCTTTCTTCACTTAAGCCGCTCA 1560
1501 tgtgattaacatgctgaccaaaatttaaccccaagaagctgaaccttcttcaactlaagccgcaa 1560
1561 TTGGAACCTTGGACCAAGTGCGCGAGTGCTCAGCTTGCGACTTCTGTCCACACCAAGC 1620
1561 ttggaaccttggaccaaagtgcgagtgctcagctggcagcttctgtccacaccaaagc 1620
1621 GAGGCTGAGCATGCGAGCAGCTGACAAAGCCTGGCTGAGAAAGCTCTAAGGCTGTGTGA 1680
1621 gaggctgagcatgcgagcagctgacaaagcctggctgagaaagctcttaggtgtga 1680
1681 ACTACTAGGGGTGTGACATGACATGAGCTTAAATTTGCAAGAAAGAACATGCTGTGCAAG 1740
1681 actactaggggtgtgacatgacatgagctttaaatttgcagaaagaaatagctgtgcaag 1740
1741 GCTTCTGCTTCTGGGCTGCTGACAAATATGATCGACTTGTGGAATAATATATATCTTG 1800
1741 gcttctgcttctgggctgctgacaaatgatcgacttgtggaataatataatcttg 1800
1801 CCTTTTGAATGAAGGCTTACATGAGGTTTCATCAGCAAGAGCGGAGCGGCATTC 1860
1801 ccttttgaatgaaggcttacctgaggtttcatcagcaagagcgagcgagcgccatcc 1860
1861 TAAACCAAAAGCCCCCGCGCACCTTCTACGTGGCTTCACGAGACAGCAAGAAAGAG 1920
1861 taaccaaagcccccgcgcaccttctacgtggcttcacgagagagcagaagaagag 1920
1921 GGGTCACCTTACTTGTGGTGGAAGAGACATGAGTGCAGACCCAGATCCACTTCTAG 1980
1921 gggtcaccttacttgtggtggaagagacatgagtgccagacccagatccacttctag 1980
1981 AGCCATACACCAGCAGCAGCTGAGCAACATGTGATTTGCTGGAATGATATGCGCTATA 2040
1981 agccataccagcagcagcagctgagcaacatgtgatttgcctggaatgatatgcgctata 2040
2041 AGATCATGATGCGACCAACATCCGTTGCTGCTCAGTTGCTTACCTTACCCGACATTC 2100
2041 agatcatgatgcgaccacacatccggttgcctcagttgcttaccttaccttaccggacatcc 2100
2101 CCAAGGAGAGGCAATTTGGAAGTACTGTAGGCCCGAGAGCGCACAGCACCCGAGCCG 2160
2101 ccaaggagagcaatttggaaagtactgtagggccgagagcgaagagacacccgagcg 2160
2161 ACCCAAGTAAAGTGTCCCGGACCTGGAAGAGCAGGTTGATGTGTGACACCAAGCAGCT 2220
2161 acccaagttaaagtgtcccgacctggaagagcaggttcatctgtgtgacaccaaagcct 2220
2221 GCAGCAATACCAATTTGACCTGCGGATGTCGCCCGACACTTTAGATTTCATTTGATGCAATTTG 2280
2221 gcagcaataccatTTTgacctgCGGATGTCGCCCGACACTTTAGATTTCATTTGATGCAATTTG 2280
2281 GAAATTAACGGTGAAGTCTGAGCCCTTACAGAGAGGCAATTTGAGTGCCTGACGTTTG 2340
2281 gaaatlaacgggtgaagtctgagcccttaagcagagagcgagcttggagctgcctcaagcttg 2340
2341 ACATGATCTGACCTCGGAGTGTGCTACCTCCCATGTGAGAGCTGAACCGAAGCT 2400

|||||
Db 2341 aatgatgacgacctgagatgctgctacccctccatgagagagctgaaccagaagct 2400
Oy 2401 GCAGAGAGCTGACTTGAGACACCTGCCCCGTGCTCCACCCCTAAGACGCGGAACCCATA 2460
Db 2401 gcaagagagctgactgagagacacctgccccgctgctccacccctaaagcagcgaacccata 2460
Oy 2461 TCGTGTGAACCTCTTAACCTTTGTGCTCCAGATTTTTTTTTTAAATTTCTACTTGTCT 2520
Db 2461 tgcgtcgaacaccttaactcttggtccagatcttttttcttaattcttaactctgct 2520
Oy 2521 ATCTTGGGCATCTGCGACCTTTTAAAGAGAGAAATGATGACGTGCGGTGATGAAC 2580
Db 2521 acccttgagcaactcgggcaactctttaaagagaaatgagctgagctggggtgaataac 2580
Oy 2581 TGTATGTAAAGAGAGAGACCTGTGAGTCTGSGAGTGGGCTGAGAGCAGAGAGAGGC 2640
Db 2581 tgtatgtaagaagagagagacctctgagctctgggagctgggagctgagagagagagagc 2640
Oy 2641 AAAGGGGAACACCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700
Db 2641 aaaggggaacacctctctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2700
Oy 2701 TGTGTAGACAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760
Db 2701 tgtgtagaagaagtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 2760
Oy 2761 ACCCATCTCTGGGAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2820
Db 2761 acccatctctgggaacacctctgctgctgctgctgctgctgctgctgctgctgctgctg 2820
Oy 2821 TGGAGCTAAGCCAGAGGTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2869
Db 2821 tagagactaagccagagaggtctctctcttaataataataataataataataataata 2869

RESULT 10
US-08-416-581B-7
Sequence 7, Application US/08416581B
Patent No. 5719042
GENERAL INFORMATION:
APPLICANT: Kishimoto, Tadamiatsu
APPLICANT: Akira, Shizuo
TITLE OF INVENTION: TRANSCRIPTION FACTOR APFR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MIOM, ZINN, MACPHEK & SEARS
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416,581B
FILING DATE: 04-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-65825/1994
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Nakamura, Dean H
REGISTRATION NUMBER: 33,981
REFERENCE/DOCKET NUMBER: 0-37891
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-7860
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:

LENGTH: 2652 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
US-08-416-581B-7

Query Match 84.2% Score 2415.6 DB 1: Length 2652;
Best Local Similarity 99.3% Pred. No. 0:
Matches 2436; Conservative 0; Mismatches 14; Indels 2; Gaps 1:

Oy 13 CAGCCGCGCCAGTTCGGGC--TCAGCCCGAGACAGTTCGAGACCCCTGACTGCAAGAGAT 70
Db 201 CAGCCGCGCCAGTTCGGGCCTCAGCCCGGAGAGAGTTCGAGACCCCTGACTGCAAGAGAT 260
Oy 71 GGCCTAGTGAACACACTGACAGCTGAGACACAGCTTACCTGAAACAGCTGACACAGCT 130
Db 261 GGCCTAGTGAACACACTGACAGCTGAGACACAGCTTACCTGAAACAGCTGACACAGCT 320
Oy 131 GTACATCGACAGCTTCCCATGAGAGCTGGGAGCTTCTGACACCTTGTGATTGAGATGA 190
Db 321 GTACATCGACAGCTTCCCATGAGAGCTGGGAGCTTCTGACACCTTGTGATTGAGATGA 380
Oy 191 AGACTGGCATATGCAAGCCAGCAAGAGATGACATGCGACGTTGGTGTTCATATATCTTT 250
Db 381 AGACTGGCATATGCAAGCCAGCAAGAGATGACATGCGACGTTGGTGTTCATATATCTTT 440
Oy 251 GGGTGAATTTGACACAGCATATATAGCCGATTCCTGCAAGAGTCCAAATGTCCTATGACA 310
Db 441 GGGTGAATTTGACACAGCATATATAGCCGATTCCTGCAAGAGTCCAAATGTCCTATGACA 500
Oy 311 CAACCTTGAAAGATCAAGACAGTTCCTGCAAGAGATTCCTGCAAGAGTTCCTGCAAGAG 370
Db 501 CAACCTTGAAAGATCAAGACAGTTCCTGCAAGAGATTCCTGCAAGAGTTCCTGCAAGAG 560
Oy 371 TGCCCGGATCTGGGCGGCGGATGCGGTGGGAAGATCTGCGCTCTCCACAGGCGACCCAC 430
Db 561 TGCCCGGATCTGGGCGGCGGATGCGGTGGGAAGATCTGCGCTCTCCACAGGCGACCCAC 620
Oy 431 GGCAGGCCAGCAAGGGGGGCGAGGCCAACCACCCCAACAGCGCGGTAGTGAAGAGAGA 490
Db 621 GGCAGGCCAGCAAGGGGGGCGAGGCCAACCACCCCAACAGCGCGGTAGTGAAGAGAGA 680
Oy 491 GCAGATGTTGGAGACACATCTTTCAGATGTCGGGAAGCAAGTGCAGAGATCTAGACAGA 550
Db 681 GCAGATGTTGGAGACACATCTTTCAGATGTCGGGAAGCAAGTGCAGAGATCTAGACAGA 740
Oy 551 AATGAAGTGTGGAGACACCTCCAGAGACAGCTTGTTCAACTACAAACCCCTCAAGAG 610
Db 741 AATGAAGTGTGGAGACACCTCCAGAGACAGCTTGTTCAACTACAAACCCCTCAAGAG 800
Oy 611 CCAGGAGACATGACAGATCTGAATGAAACACAGCTTGTGACCAAGACAGAGATGCA 670
Db 801 CCAGGAGACATGACAGATCTGAATGAAACACAGCTTGTGACCAAGACAGAGATGCA 860
Oy 671 GCAGTGGAAAGATGCTCACAGCCCTGACAGAGAGGCGGAAGAGATGTGATGATGCT 730
Db 861 GCAGTGGAAAGATGCTCACAGCCCTGACAGAGAGGCGGAAGAGATGTGATGATGCT 920
Oy 731 GGGGGGGCTTTGTGCAAGTGGAGTGTGAGAGACACTGACTGATGAAGAGTGGC 790
Db 921 GGGGGGGCTTTGTGCAAGTGGAGTGTGAGAGACACTGACTGATGAAGAGTGGC 980
Oy 791 TGACTGGAAGAGGGCGCCAGAGATGCGGTGATGGAGAGCCCTCCCAACATGCTGCTGA 850
Db 981 TGACTGGAAGAGGGCGCCAGAGATGCGGTGATGGAGAGCCCTCCCAACATGCTGCTGA 1040
Oy 851 CCGTGTGAAATGATGATCTTATTTAGCAATCTCAACTTGAAGACCCGCAACAAAT 910
Db 1041 CCGTGTGAAATGATGATCTTATTTAGCAATCTCAACTTGAAGACCCGCAACAAAT 1100
Oy 911 TAAGAAACTGAGAGAGCTGCACAGCAAAAGTGTCTTCAAGAGGCGAGCCCTATGCTGCAGA 970

|||||
Db 1101 TAGAAACTGGAGAGCTGACGAGAAAGTGTCTTACAGGGCCACCTATCTGACACA 1160
Qy 971 CCGGCCATGCTGGAGAGAGAGATGCTGAGCTGTTGAGAACTTAATGAAGATGTGCTT 1030
Db 1161 CCGGCCATGCTGGAGAGAGAGATGCTGAGCTGTTGAGAACTTAATGAAGATGTGCTT 1220
Qy 1031 CGTGGTGGAGCGGACGCTGCTGATGCCCACCGGACCGGCTTACTGATCAAGAC 1090
Db 1221 CGTGGTGGAGCGGACGCTGCTGATGCCCACCGGACCGGCTTACTGATCAAGAC 1280
Qy 1091 TGGTGTGCTGATTCACAGAAAGTCAGTTGCTGTCAAATTTCTGAGTTGAATATCA 1150
Db 1281 TGGTGTGCTGATTCACAGAAAGTCAGTTGCTGTCAAATTTCTGAGTTGAATATCA 1340
Qy 1151 GCTTAAATTAAGTGTGCTGATTAAGACTGTGGGATGTCTGCTCCCTCAGAGGTC 1210
Db 1341 GCTTAAATTAAGTGTGCTGATTAAGACTGTGGGATGTCTGCTCCCTCAGAGGTC 1400
Qy 1211 TCGGAATTTAACATTTCTGGGACGAGACACAAAGTGATGAGATGAGAGTGAACAA 1270
Db 1401 TCGGAATTTAACATTTCTGGGACGAGACACAAAGTGATGAGATGAGAGTGAACAA 1460
Qy 1271 CGGAGCTGTCTGAGAGTTCAAGACCTGACCTTAGGGAGCAGAGATGTGGGAATGG 1330
Db 1461 CGGAGCTGTCTGAGAGTTCAAGACCTGACCTTAGGGAGCAGAGATGTGGGAATGG 1520
Qy 1331 AGGCGGTGCCAATTTGTATGCTCTGATGATGCTGAGTGGAGGCTGACCTGATCACCTT 1390
Db 1521 AGGCGGTGCCAATTTGTATGCTCTGATGATGCTGAGTGGAGGCTGACCTGATCACCTT 1580
Qy 1391 CGAGACTGAGGTGTACGACCAAGGCTTCAGATTGACCTAGACCACTCTTGCCAGT 1450
Db 1581 CGAGACTGAGGTGTACGACCAAGGCTTCAGATTGACCTAGACCACTCTTGCCAGT 1640
Qy 1451 TGTGTGATCTCCCAACATCTGTGATGCCAAATGCTTGGGATCAATCTCTGTGATAA 1510
Db 1641 TGTGTGATCTCCCAACATCTGTGATGCCAAATGCTTGGGATCAATCTCTGTGATAA 1700
Qy 1511 CATGCTGACCAATTAACCCCAAGAACTGTAATCTTCACTAAGCGGCAATTTGGAACCTG 1570
Db 1701 CATGCTGACCAATTAACCCCAAGAACTGTAATCTTCACTAAGCGGCAATTTGGAACCTG 1760
Qy 1571 GGACCAAGTGGCGGAGTGTCTGACCTGGAGTTCCTGCCACCAAGCGGAGCTGAG 1630
Db 1761 GGACCAAGTGGCGGAGTGTCTGACCTGGAGTTCCTGCCACCAAGCGGAGCTGAG 1820
Qy 1631 CATGAGCAGCTGACCAAGCGTGGCTGAGAAAGTCTCTAGGGCTGTGTAATCTACTAGG 1690
Db 1821 CATGAGCAGCTGACCAAGCGTGGCTGAGAAAGTCTCTAGGGCTGTGTAATCTACTAGG 1880
Qy 1691 GTGTGATGATCATGAGCTTAATTTCTGCAAAAGAAACATGGCTGGCAAGGGCTTCTCTT 1750
Db 1881 GTGTGATGATCATGAGCTTAATTTCTGCAAAAGAAACATGGCTGGCAAGGGCTTCTCTT 1940
Qy 1751 CTGGCTGTGGCTAGCAATATCATGACCTGTGAAAAAGATATCTTGGCCCTTTGGAA 1810
Db 1941 CTGGCTGTGGCTAGCAATATCATGACCTGTGAAAAAGATATCTTGGCCCTTTGGAA 2000
Qy 1811 TGAAGGTATCATGAGTGTTCATCAGCAAGGAGCGGAGCGGCAATCTTAAGCACAAA 1870
Db 2001 TGAAGGTATCATGAGTGTTCATCAGCAAGGAGCGGAGCGGCAATCTTAAGCACAAA 2060
Qy 1871 GCGCCCGGACCTTCTCTACTGCGCTTACGGAAGACGCAAGAGAGAGGGGTCTACTTT 1930
Db 2061 GCGCCCGGACCTTCTCTACTGCGCTTACGGAAGACGCAAGAGAGAGGGGTCTACTTT 2120
Qy 1931 CACTTGGGTGAAAAAGACATCACTGGCAAGACCAAGATCCAGTCTAGAGCCATACAC 1990
Db 2121 CACTTGGGTGAAAAAGACATCACTGGCAAGACCAAGATCCAGTCTAGAGCCATACAC 2180
Qy 1991 CAAGCAGCAGCTGAACAACATGTCTTGTGAATCATGAGGCTATAAGATCATGCA 2050
|||||

Db 2181 CAAGCAGCAGCTGAACAACATGTCTATTTGCTGAATCATCTAGGCTATAAGATCATGGA 2240
Qy 2051 TGCAGCAACATCTGCTGTCTGCTCAGTTGTCTACTTACCCGACATTTCCCAAGAGGA 2110
Db 2241 TGCAGCAACATCTGCTGTCTGCTCAGTTGTCTACTTACCCGACATTTCCCAAGAGGA 2300
Qy 2111 GGCAATTTGGAAGTACTGTAGCCCGGAGAGAGCAGACAGACCCCAAGCCGAGCCAGGTAG 2170
Db 2301 GGCAATTTGGAAGTACTGTAGCCCGGAGAGAGCAGACAGACCCCAAGCCGAGCCAGGTAG 2360
Qy 2171 TGTGTGCCCCGTACCTGGAAGACCAAGTTTCATCTGTGTGACACCAAGACCTGACCAATAC 2230
Db 2361 TGTGTGCCCCGTACCTGGAAGACCAAGTTTCATCTGTGTGACACCAAGACCTGACCAATAC 2420
Qy 2231 CATTGACCTGCGGATGTCTCCCGGACCTTTAGATTCAATTGATGAGTTGGAATAACCG 2290
Db 2421 CATTGACCTGCGGATGTCTCCCGGACCTTTAGATTCAATTGATGAGTTGGAATAACCG 2480
Qy 2291 TGAAGTGTGAGGCTGACGAGGAGGAGGAGGAGGAGTTGAGTGTGAGTGTGACATGATCT 2350
Db 2481 TGAAGTGTGAGGCTGACGAGGAGGAGGAGGAGGAGTTGAGTGTGAGTGTGACATGATCT 2540
Qy 2351 GACCTGGAGTGTGCTTACCTCCCATGTGAGAGCTGGAACCAAGAGCTGACAGACGT 2410
Db 2541 GACCTGGAGTGTGCTTACCTCCCATGTGAGAGCTGGAACCAAGAGCTGACAGACGT 2600
Qy 2411 GACTTGAGACACATGCCCCGCTGCTCACCCCTTAAGCAGCAGCAACCCCATATC 2462
Db 2601 GACTTGAGACACATGCCCCGCTGCTCACCCCTTAAGCAGCAGCAACCCCATATC 2652
RESULT 11
US-08-416-581B-8
Sequence 8, Application US/08416581B
Patent No. 5719042
GENERAL INFORMATION:
APPLICANT: Kishimoto, Tadimitsu
APPLICANT: Akira, Shizuo
TITLE OF INVENTION: TRANSCRIPTION FACTOR APRE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MION, ZINN, MACPHEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416,581B
FILING DATE: 04-Apr-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-65825/1994
FILING DATE: 04-Apr-1994
ATTORNEY/AGENT INFORMATION:
NAME: Nakamura, Dean H.
REGISTRATION NUMBER: 33,981
REFERENCE/DOCKET NUMBER: Q-37891
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-7860
INFORMATION FOR SEQ. ID NO.: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2652 base pairs
type: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA

ORIGINAL SOURCE:
ORGANISM: Mouse
TISSUE TYPE: Liver
FEATURE:
NAME/KEY: CDS
LOCATION: 259..2568
US-08-416-5818-8

Query Match 84.2% Score 2415.6; DB 1; Length 2652;
Best Local Similarity 99.3% Pred. No. 0;
Matches 2436; Conservative 0; Mismatches 14; Indels 2; Gaps 1;

QY 13 CAGCCGGCCACTCGGCGC--TCAGCCCGAGACACTCGAGACCCCTGACTCGACAGAGAT 70
DB 201 CAGCCGGCCACTCGGCGCTCGAGCCGAGACACTCGAGACCCCTGACTCGACAGAGAT 260
QY 71 GGCTAGTGGAAACACAGCTGACAGCTGACACAGCTACCTGGAAGCAGCTGCACAGCT 130
DB 261 GGCTAGTGGAAACACAGCTGACAGCTGACACAGCTACCTGGAAGCAGCTGCACAGCT 320
QY 131 GACACACGACGTTCCCATGAGCTCGGAGTTCTGACACCTTGAGTGGAGATCA 190
DB 321 GTACAGCCGACACTTCCCATGAGCTCGGAGTTCTGACACCTTGAGTGGAGATCA 380
QY 191 AGACTGGGCATATGACGACGACCAAGAGTACATGCCACGTTGGTGTTCATATCTCT 250
DB 381 AGACTGGGCATATGACGACGACCAAGAGTACATGCCACGTTGGTGTTCATATCTCT 440
QY 251 GGSTGAATGACGACCAATATAGCCGATCTCGCAAGAGTCCATGTCCTATACAGA 310
DB 441 GGSTGAATGACGACCAATATAGCCGATCTCGCAAGAGTCCATGTCCTATACAGA 500
QY 311 CAACCTTGCAGCAATCAAGCAAGTCTGACAGCAGTATCTTGAAGACCAATGGAAT 370
DB 501 CAACCTTGCAGCAATCAAGCAAGTCTGACAGCAGTATCTTGAAGACCAATGGAAT 560
QY 371 TGCCTGGATGCTGGGCGGCGATGCTGGGGAAGATCTGCCCTGACAGAGGCGAC 430
DB 561 TGCCTGGATGCTGGGCGGCGATGCTGGGGAAGATCTGCCCTGACAGAGGCGAC 620
QY 431 GGCAGCCGACGAAGGGGGCAGGCCAACCCACAGCCGCTGATGACAGAGA 490
DB 621 GGCAGCCGACGAAGGGGGCAGGCCAACCCACAGCCGCTGATGACAGAGA 680
QY 491 GGAGATGTTGGAGCAGCATCTCAGAGATGTCGGAAGCGAGTGCAGAGATCTGAACAGA 550
DB 681 GGAGATGTTGGAGCAGCATCTCAGAGATGTCGGAAGCGAGTGCAGAGATCTGAACAGA 740
QY 551 AATGAAGTGTGTGAGAACTCCAGAGACACTTGTATTTCACTACAAAACCTCAAGAG 610
DB 741 AATGAAGTGTGTGAGAACTCCAGAGACACTTGTATTTCACTACAAAACCTCAAGAG 800
QY 611 CCAAGGAGACATGACAGATCTGAATGGAACAACAGTGTGAGACGACAGCAAGATGCA 670
DB 801 CCAAGGAGACATGACAGATCTGAATGGAACAACAGTGTGAGACGACAGCAAGATGCA 860
QY 671 GCAGCTGGAACAGATGCTCACAGCCCTGAGACAGATGCGAGAAGCAGTGTGAGTGA 730
DB 861 GCAGCTGGAACAGATGCTCACAGCCCTGAGACAGATGCGAGAAGCAGTGTGAGTGA 920
QY 731 GCGGGGGCTCTTGTGACAGCAATGAGTACGTGACAGAAACACTGACTGATGAAGAGCTGC 790
DB 921 GCGGGGGCTCTTGTGACAGCAATGAGTACGTGACAGAAACACTGACTGATGAAGAGCTGC 980
QY 791 TGACGTGGAAGAGGGCGGACAGATGCGGTGACAGAGGCGCTCCCAACATGTGCTTGA 850
DB 981 TGACGTGGAAGAGGGCGGACAGATGCGGTGACAGAGGCGCTCCCAACATGTGCTTGA 1040
QY 851 CCGTCTGAAAACTGGATTAATTCATTACAGCAATCTCAACTTCAGACCCGCCCAACAAT 910
DB 1041 CCGTCTGAAAACTGGATTAATTCATTACAGCAATCTCAACTTCAGACCCGCCCAACAAT 1100

QY 911 TAAGAACTGAGAGAGCTGCAGCAGAAAGTCTCTACAAAGGGCAGCCCTATCTGCAGCA 970
DB 1101 TAAGAACTGAGAGAGCTGCAGCAGAAAGTCTCTACAAAGGGCAGCCCTATCTGCAGCA 1160
QY 971 CCGGCCCATGCTGAGAGAGAGATCTGTGAGCTCTTCCAGAACTTAATGAAGAGTGCCT 1030
DB 1161 CCGGCCCATGCTGAGAGAGAGATCTGTGAGCTCTTCCAGAACTTAATGAAGAGTGCCT 1220
QY 1031 CGTGTGAGAGCGGCGAGCCCTGATGCCCCATGACCCCGAGCCGCGCTTACTCATCAAGAC 1090
DB 1221 CGTGTGAGAGCGGCGAGCCCTGATGCCCCATGACCCCGAGCCGCGCTTACTCATCAAGAC 1280
QY 1091 TGTGTCCAGTTTACACAGAAAGTACGTGCTGTCATAATTTCTGTAGTTGAATATCA 1150
DB 1281 TGTGTCCAGTTTACACAGAAAGTACGTGCTGTCATAATTTCTGTAGTTGAATATCA 1340
QY 1151 GCTTAAATTTAAAGTGTGATTAAGAGCTCTGGGAGATTTGCTGCCCTCAGAGGCTC 1210
DB 1341 GCTTAAATTTAAAGTGTGATTAAGAGCTCTGGGAGATTTGCTGCCCTCAGAGGCTC 1400
QY 1211 TCGGAATTTTAACATTTGCGGCACGACACAAAGAGTGAACATGAGAGCTTAACAA 1270
DB 1401 TCGGAATTTTAACATTTGCGGCACGACACAAAGAGTGAATTAACATGAGAGCTTAACAA 1460
QY 1271 CGGACGCTGTCTGAGAGTTCAACGACCTGACCTTAGGAGAGAGATGAGGAATGG 1330
DB 1461 CGGACGCTGTCTGAGAGTTCAACGACCTGACCTTAGGAGAGAGATGAGGAATGG 1520
QY 1331 AGGCCGTGCAATTTGTGATGCTCTCTTATGCTGAGTACTGAGAGCTGCACCTGATCACCT 1390
DB 1521 AGGCCGTGCAATTTGTGATGCTCTCTTATGCTGAGTACTGAGAGCTGCACCTGATCACCT 1580
QY 1391 CGAGCTGAGGTGTACACACCAAGGCTCAAGATGACTTAAGAGACCCACTCTTGCCAGT 1450
DB 1581 CGAGCTGAGGTGTGTACACCAAGGCTCAAGATGACTTAAGAGACCCACTCTTGCCAGT 1640
QY 1451 TGTGTGATCTCCACATCTGTCAAGTGCCTGAGTGCCTGAGCATCAATCTGTGATATA 1510
DB 1641 TGTGTGATCTCCACATCTGTCAAGTGCCTGAGTGCCTGAGCATCAATCTGTGATATA 1700
QY 1511 CATGCTGACCAATTAACCCCAAGACGTAAGTCTTCACTTAAGCCGCCAATTTGGAACCTG 1570
DB 1701 CATGCTGACCAATTAACCCCAAGACGTAAGTCTTCACTTAAGCCGCCAATTTGGAACCTG 1760
QY 1571 GGACCAAGTGGCGAGGTGCTGAGTGCAGTTCGTCACACACCAAGCAAGCAAGCTGAG 1630
DB 1761 GGACCAAGTGGCGAGGTGCTGAGTGCAGTTCGTCACACACCAAGCAAGCAAGCTGAG 1820
QY 1631 CATGCAATTTTGTGACACGCTGCTGAGAGCTCTAGAGCCCTGTGTGAACATCTCAGG 1690
DB 1821 CATGCAATTTTGTGACACGCTGCTGAGAGCTCTAGAGCCCTGTGTGAACATCTCAGG 1880
QY 1691 GTGTGATGATCATGAGGCTTAATTTCTGCAAGAAACATGCTGCAAGGGCTCTCTCT 1750
DB 1881 GTGTGATGATCATGAGGCTTAATTTCTGCAAGAAACATGCTGCAAGGGCTCTCTCT 1940
QY 1751 CTGGCTGCTGAGCAATATCATGACCTGTGTGAAAAGTATATCTTGGCCCTTTGAA 1810
DB 1941 CTGGCTGCTGAGCAATATCATGACCTGTGTGAAAAGTATATCTTGGCCCTTTGAA 2000
QY 1811 TGAAGGTAACATCATGAGGTTTATCATGACAGAGAGCGGAGCGGCATCTTAAGCACAA 1870
DB 2001 TGAAGGTAACATCATGAGGTTTATCATGACAGAGAGCGGAGCGGCATCTTAAGCACAA 2060
QY 1871 GCGCCCGGACCTTCTACTGCGCTTACAGCAGAGAGCAAGAAAGAGAGGGGTACCTT 1930
DB 2061 GCGCCCGGACCTTCTACTGCGCTTACAGCAGAGAGCAAGAAAGAGAGGGGTACCTT 2120
QY 1931 CACTTGGGTGAAAAGACATCATGTGCAAGACCCAGATCCAGTCTGTGAAGCAATACAC 1990
DB 2121 CACT TGGGTGAAAAGACATCATGTGCAAGACCCAGATCCAGTCTGTGAAGCAATACAC 2180
QY 1991 CAAGCAAGACGTGAACACATGTCTTTGTGAAATCATCATGGGCTTAAGATCATGGA 2050


```

Db 961 TTTCTGTGGGCGGAGCCCTGCATGCCATGCCGCGGCCCTTAGTATCAAG 1020
Oy 1089 ACTGCTGCCAGTTTACACGAAGTCAGTGTGTCATTAATTCTCGAGTTGAATTA 1148
Db 1021 ACTGCTGCCAGTTTACACGAAGTCAGTGTGTCATTAATTCTCGAGTTGAATTA 1080
Oy 1149 CAGCTTAAATTAAGTGTGATTAAGATCTCTGGGATGTTGCTGCTCAGAGGG 1208
Db 1081 CAGCTTAAATTAAGTGTGATTAAGATCTCTGGGATGTTGCTGCTCAGAGGG 1140
Oy 1209 TCTCGAAATTAATTAATTCCTGGGACGACAAAGATGATGAGAGAGCTTAAC 1268
Db 1141 TCTCGAAATTAATTAATTCCTGGGACGACAAAGATGATGAGAGAGCTTAAC 1200
Oy 1269 AACGCGACCCGTCTGCGAGATTCAGACCCCTGACCCCTTAGAGAGCAGATGTGGAA 1338
Db 1201 AACGCGACCCGTCTGCGAGATTCAGACCCCTGACCCCTTAGAGAGCAGATGTGGAA 1260
Oy 1329 GGAGGCCGTGCCAATTTGTGATGCTCTTGTATCGTATGAGAGCTGCACCTGAC 1388
Db 1261 GGAGGCCGTGCCAATTTGTGATGCTCTTGTATCGTATGAGAGCTGCACCTGAC 1320
Oy 1389 TTCGAGACTGAGTGTACACCAAGGCTCAAGATTGACCTAGAGACCCACCTTGCCA 1448
Db 1321 TTCGAGACTGAGTGTACACCAAGGCTCAAGATTGACCTAGAGACCCACCTTGCCA 1380
Oy 1449 GTTGTGTGATCTCCACATCTGTACATGCCAAATGCTTGGCATCAATCTGTGAT 1508
Db 1381 GTTGTGTGATCTCCACATCTGTACATGCCAAATGCTTGGCATCAATCTGTGAT 1440
Oy 1509 AACATGCTGACCAATTAACCCCAAGACGTGAACCTTCTCAATAAGCCGCAATTGGAAC 1568
Db 1441 AACATGCTGACCAATTAACCCCAAGACGTGAACCTTCTCAATAAGCCGCAATTGGAAC 1500
Oy 1569 TGGGACCAAGTGGCCGAGGTCTCAGCTGGCAGTTCTGTCACACCACCAAGCGAGCTG 1628
Db 1501 TGGGACCAAGTGGCCGAGGTCTCAGCTGGCAGTTCTGTCACACCACCAAGCGAGCTG 1560
Oy 1629 AGCATCGAGCTGACCAAGCCTGGCTGAGAAGCTCTAGAGGCTGGGTGTAAGTACTCA 1688
Db 1561 AGCATCGAGCTGACCAAGCCTGGCTGAGAAGCTCTAGAGGCTGGGTGTAAGTACTCA 1620
Oy 1689 GGGTGTGATTCACATGGGCTAAATTTCTGCAAGAAAACATGGCTGGCAAGGGCTTCTCC 1748
Db 1621 GGGTGTGATTCACATGGGCTAAATTTCTGCAAGAAAACATGGCTGGCAAGGGCTTCTCC 1680
Oy 1749 TTTGGGCTGTGGCTAGACAATATCATGCACTTGTGAAGAAAGTATCTTGGCCCTTGG 1808
Db 1681 TTTGGGCTGTGGCTAGACAATATCATGCACTTGTGAAGAAAGTATCTTGGCCCTTGG 1740
Oy 1809 AATGAAGGTCATCATGAGTTTCATCAGCAAGAGCGGAGCGGCCATCTTAAGCACA 1868
Db 1741 AATGAAGGTCATCATGAGTTTCATCAGCAAGAGCGGAGCGGCCATCTTAAGCACA 1800
Oy 1869 AAGCCCCCGGACCTTCTCTACTGCGCTTCAGCGAGACGACAAAGAGAGGGTCACT 1928
Db 1801 AAGCCCCCGGACCTTCTCTACTGCGCTTCAGCGAGACGACAAAGAGAGGGTCACT 1860
Oy 1929 TTTCACTTGGTGGAAGAGATCAGTGGCAAGACCCAGATCCAGTCTGTAGAGCAATAC 1988
Db 1861 TTTCACTTGGTGGAAGAGATCAGTGGCAAGACCCAGATCCAGTCTGTAGAGCAATAC 1920
Oy 1989 ACCAAGCAGCAGCTGAACACATGCTATTCCTGAATATCATATGAGCTATTAAGATCATG 2048
Db 1921 ACCAAGCAGCAGCTGAACACATGCTATTCCTGAATATCATATGAGCTATTAAGATCATG 1980
Oy 2049 GATGCGACCAACATCTGCTGCTCAGTCTTCACTTCACTCCCGGATATCCCAAGAG 2108
Db 1981 GATGCGACCAACATCTGCTGCTCAGTCTTCACTTCACTCCCGGATATCCCAAGAG 2040
Oy 2109 GAGGCAATTTGGAAGTACTTAGCGCCGAGAGCCAGAGACCCCGAAGCCGACCAAGT 2168

```

```

Db 2041 GAGGCAATTTGGAAGTACTTAGCGCCGAGAGCCAGACACCCCGAAGCCGACCAAGT 2100
Oy 2169 AGTCTGCCCGCTACCTGAAGACCAAGTTCAATCTGTGTGACAGCAGACCTGACGAAT 2228
Db 2101 AGTCTGCCCGCTACCTGAAGACCAAGTTCAATCTGTGTGACAGCAGACCTGACGAAT 2160
Oy 2229 ACCATTTGACCTGCGGATGTCTCCCGGACACTTACATTTGATGATGAGTGGAAATAC 2288
Db 2161 ACCATTTGACCTGCGGATGTCTCCCGGACACTTACATTTGATGATGAGTGGAAATAC 2220
Oy 2289 GGTGAAGTGTGAGCCCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2348
Db 2221 GGTGAAGTGTGAGCCCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2280
Oy 2349 CTGACCTGAGAGTGTGCTACCTCCCGCATG 2378
Db 2281 CTGACCTGAGAGTGTGCTACCTCCCGCATG 2310

RESULT 13
US-08-416-581B-3
: Sequence 3, Application US/08416581B
: Patent No. 5719042
: GENERAL INFORMATION:
: APPLICANT: Kishimoto, Tadamiatsu
: APPLICANT: Akira, Shizuo
: TITLE OF INVENTION: TRANSCRIPTION FACTOR APF
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SUGHRUE, MION, ZINN, MACBEAR & SEAS
: STREET: 2100 Pennsylvania Avenue, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/416,581B
: FILING DATE: 04-Apr-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 6-65825/1994
: FILING DATE: 04-Apr-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Nakamura, Dean H.
: REGISTRATION NUMBER: 33,981
: REFERENCE/DOCKET NUMBER: Q-37891
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)293-7060
: TELEFAX: (202)293-7860
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2787 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: US-08-416-581B-3

Query Match 74.7%; Score 2144.2; DB 1; Length 2787;
Best Local Similarity 90.9%; Pred. No. 0;
Matches 2305; Conservative 0; Mismatches 223; Indels 7; Gaps 2;
Oy 61 GCACGAGATGGCTCAGTGAACCAAGCTGACAGCTGAGACACAGCTACCTGAAGCAGC 120
Db 213 GAAACAGGATGGCCCAATGGAATCAGCTACAGCAGAGCTTGACACACGAGTACCTGGAGCAGC 272
Oy 121 TGCAACAGCTGTACAGGACACAGCTTCCCATGAGAGCTGCGGCAATTCCTGACACTTGA 180

```


||||| 1233 TCATCAAGACCGCGCTCCAGTCTACTACATAAATCAGGTTGCGTCAAGTCCCTGAGT 1292
1141 TGAATATATAGCTAAATTAATTAAGTGTGATTAAGACTCTGGGATGTTGCTGCC 1200
1293 TGAATATATAGCTAAATTAATTAAGTGTGATTAAGACTCTGGGATGTTGCTGCC 1352
1201 TCAGAGGCTCTCGAAATTTACATTTCTGGGCGAACAACAAAGTATGATGATGAGG 1260
1353 TCAGAGGATCCCGGAATTTACATTTCTGGGCGAACAACAAAGTATGATGATGAGG 1412
1261 AGCTCAACAAGGCGCTCTGAGAGTTCAGGAGTCCAGCCCTTAGGAGCAGAGAT 1320
1413 AATTCACAAGGCGCTCTGAGAGTTCAGGAGTTCAGGAGTTCAGGAGCAGAGAT 1472
1321 GTGGAGATGAGGCGCTGCGCAATTTGATGCTCTTGTATGCTGACTGAGAGCTGCACC 1380
1473 GTGGAGATGAGGCGCGAGCAATTTGATGCTCTTGTATGCTGACTGAGAGCTGCACC 1532
1381 TGATACCTTCGAGACTGAGGTATACCAAGCCCAAGGCTCAAGATTGACTGAGAGCCACT 1440
1533 TGATACCTTCGAGACTGAGGTATACCAAGGCTCAAGATTGACTGAGAGCCACT 1592
1441 CCTGCGAGTTGATGATGCTCAACATCTGTAGATGCCAAATGCTTGGGATCAATCC 1500
1593 CCTTGTACGTTGATGATGCTCAACATCTGTAGATGCCAAATGCTTGGGATCAATCC 1652
1501 TGTGATATAACATGCTGAGCAATTAACCCCAAGACGTAACCTTCTTCAAGCCGCCAA 1560
1653 TGTGATATAACATGCTGAGCAATTAACCCCAAGACGTAACCTTCTTCAAGCCGCCAA 1712
1561 TTGGAACTTGGGACCAAGTGGCGGAGTGTCTCAGCTGGGAGTTCTGCTCCACACCAAGC 1620
1713 TTGGAACTTGGGACCAAGTGGCGGAGTGTCTCAGCTGGGAGTTCTGCTCCACACCAAGC 1772
1621 GAGGCGTGAACATCGAGACAGCTGACAAAGCTGCTGAGAGCTCTTGGGCTGTGTGA 1680
1773 GAGGCGTGAACATCGAGACAGCTGACAAAGCTGCTGAGAGCTCTTGGGCTGTGTGA 1832
1681 ACTACTGAGGCTGTAGATACATGAGGCTAACTTCTGCAAGAAACATGCTGCTGAGG 1740
1833 ACTACTGAGGCTGTAGATACATGAGGCTAACTTCTGCAAGAAACATGCTGCTGAGG 1892
1741 GCTTTCCTTCTGAGGCTGTAGATACATGAGGCTAACTTCTGCAAGAAACATGCTGAGG 1800
1893 GCTTTCCTTCTGAGGCTGTAGATACATGAGGCTAACTTCTGCAAGAAACATGCTGAGG 1952
1801 CCTTTGGAATGAAGGTTACATGAGTTTATCAGCAAGGAGCGGAGCGCCATCC 1860
1953 CCTTTGGAATGAAGGTTACATGAGTTTATCAGCAAGGAGCGGAGCGCCATCC 2012
1861 TAAGCAAAAGCCCCCGGAGCACTTCTACTGGCTTACAGGAGAGCAAGAAAGAGAG 1920
2013 TGAGCACTAAGCCCCCGGAGCACTTCTACTGGCTTACAGGAGAGCAAGAAAGAGAG 2072
1921 GGGTCACTTCTGAGTGGTGAAGAAAGACATCAGTGGCAAGACCCAGATCAGTCTAG 1980
2073 GCGTCACTTCTGAGTGGTGAAGAAAGACATCAGGCGTGAAGACCCAGATCAGTCTAG 2132
1981 AGCCATACCAAGCAGCAGCTGAACAACATGTCATTTGCTGAATCATGAGGCTATA 2040
2133 AACCATACCAAGCAGCAGCTGAACAACATGTCATTTGCTGAATCATGAGGCTATA 2192
2041 AGATCATGATGGAGCAACATCTGTGTCTCAGCTTGTACTACTTACCTTACCCGACATTC 2100
2193 AGATCATGATGGAGCAACATCTGTGTCTCAGCTTGTACTACTTATCTTATCTGACATTC 2252
2101 CCAAGGAGAGGAGCTTTGAAAGTACTAGGCGCGAGAGCAGAGACACCCGAGACCG 2160
2253 CCAAGGAGAGGAGCTTTGAAAGTACTAGGCGCGAGAGCAGAGACATCTCTGAACCTG 2312
2161 ACCGAGTACTGCTCCCTTACTGAGACCAAGTTCTGTGTGACACCAAGCACT 2220
|||||

Db 2313 ACCGAGTACGCTGCCCATACCTGAGAGCAAGTTTATCTGTGTGACACCAAGCACT 2372
Qy 2221 GCAGCAATACCATGACCTGCCGATGTCGCCCGCACCTTTAGATTGATGACGATTG 2280
Db 2373 GCAGCAATACCATGACCTGCCGATGTCGCCCGCACCTTTAGATTGATGATCAGATTG 2432
Qy 2281 GAATATACGTTGAAGTGTGAGGCTCTACGAGAGAGGAGGAGTTGATGATGATGATG 2340
Db 2433 GAATATACGTTGAAGTGTGAGGCTCTACGAGAGAGGAGGAGTTGATGATGATGATG 2492
Qy 2341 ACATGATGACCTGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2400
Db 2493 ACATGATGACCTGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2552
Qy 2401 GC---AGAGACGCTGACCTGAGACACCTGCCCGCTGTCCACCCCTTAAGCAGCAACC 2456
Db 2553 GCAGAAAGATACACCTGAGAGCCTGACCTGATTCCTGACCCCTCAGACAGCAACCC 2612
Qy 2457 CATATGCTGAAACCTCTACCTTGTGTGATGATGATGATGATGATGATGATGATGATGAT 2516
Db 2613 CAGATCATCTGAAACCTCTACCTTGTGTGATGATGATGATGATGATGATGATGATGATGAT 2669
Qy 2517 TGCATCTTTGGGCAATCTGGGACCTTTTAAAGAGAGAAATGAGTGAAGTGGGTGAT 2576
Db 2670 TGCATCTTTGGGCAATCTGGGACCTTTTAAAGATAGAGAAATGAGTGAAGTGGGTGAT 2729
Qy 2577 AACCTGTTATGTAAA 2591
Db 2730 CTGCTTTTATCTTAAA 2744

RESULT 15
US-09-288-461-1
; Sequence 1, Application US/09288461
; Patent No. 6159694
; GENERAL INFORMATION:
; APPLICANT: Karitas, James G.
; TITLE OF INVENTION: Antisense Oligonucleotide Modulation of STAT3
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPh-0338
; CURRENT APPLICATION NUMBER: US/09/288,461
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2787
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (221)-(2533)
; PUBLICATION INFORMATION:
; JOURNAL: Cell
; VOLUME: 77
; ISSUE: 1
; PAGES: 63-71
; DATE: 1994-04-08
; DATABASE ACCESSION NUMBER: L29277
; DATABASE ENTRY DATE: 1994-12-31
US-09-288-461-1

Query Match 74.7%: Score 2144.2; DB 3; Length 2787;
Best Local Similarity 90.9%: Pred. No. 0;
Matches 2305; Conservative 0; Mismatches 223; Indels 7; Gaps 2;
Qy 61 GCAGCAGAGATGCTGCTAGTGAACCGCTGAGAGCTGAGACACGCTTACTGTAAGCAGC 120
Db 213 gaacacgagatgagcccaatgaaatcagctacagcagcttgacacagctgagcagc 272
Qy 121 TGCACGAGCTGTACAGCGAGACGTTCCCATGAGAGCTGCGGAGTTCCTTGGACCTTGA 180
Db 273 tccatcagctctacagctgacagcttcccaatgagcagctgagcagcttctgagcagc 332

OY	181	TTGAGAGCTCAAGACTGGGCAATTATGACAGCCAGCAAGAGTGCACATGCCAGTTGGTGTTC	240
Db	333	ctgagaaatcaagaattctgggcaatctgagccagcaaaagaaatcacatctgaccttgggtcttc	392
OY	241	ATATATCTCTTGGGGTGAATTTGACCAGCAATATATGCGGATCTCTGGCAAGAGTCCATTGCC	300
Db	393	ataatctctctgggagagaaatttgaccagaagatatagcgcgtctcttcgaagagctcgaaatgtctc	452
OY	301	TCTATCAGCACAACTTTCGGAAGATCAAGCAGTTCCTGCAGAGCAGGTATCTTGAGAAGC	360
Db	453	ctctacagcaaatctacagaagaatcaagaagttctctcagagcagtaatctctggaagc	512
OY	361	CAATGGAATTTGGCCGGATGTGTGGCCGATACCTGTGGGAAGAAGTCTGGCCTCTCCAGA	420
Db	513	caatggagatttgcccggaattctggcccggtgctctgagggaagaatcacgcctctctaca	572
OY	421	CGCAGACCGACGGCAGCCACAGGGGGCCAGGCCAACCCCAACAGCCGCCGTATGTA	480
Db	573	ctgagccactgcggtcccgagcaaggggccaagcccaaccacgaagcagcggtggtga	632
OY	481	CAGAGAAACACACAGATTTTGAGACAGCATTTTACAGATGTCCGGAAGCAGATGCGAGATC	540
Db	633	cggagaaagcagcagatgcttgtagcagcagaccttcagatgcttcggaagagatgtagagatc	692
OY	541	TAGAAACAGAAAATTAAGGTGGTGAGAACCTCCAGAGCACATTTATTTCACTCAATAA	600
Db	693	tagaaccagaaaatgaaagtggttagagaatctccagagtgacttgattccaacttaaaa	752
OY	601	CCCTCAAGAGCCCAAGAGACATGCAGAGATCTGAATGGAANAACCAAGTCTGTGACCAGAC	660
Db	753	cccccaagagtcgaaggagacatgcaagatcttgatgtaaaacacacagtcagtgacagcgc	812
OY	661	AGAAGATGACGACCTCTGAGACACATGCTCAACCCCTGGACCAAGATGGGAGCAACATTG	720
Db	813	agaagaatgcaagcaactggaacagagatgctcaatgctctggaacagatgtagaagaaatcag	872
OY	721	TGAGTGAAGCTGGCGGGGCTTGTTCAGCAATGAGTACTGTGCAGAAAGACACTGATGTATG	780
Db	873	tgagtgagcttggtcggtggtcttctgcaagatgtagtgaagtgtagaagaaactctcacgagc	932
OY	781	AAGAGCTGTGCTGACTGGAAGAAGCGGCGACAGATGCGCTGTGATCGAGAGCCCTCCCAACA	840
Db	933	agagagcttgagcttggaataagagcgcaaaagatctgcttcgcaatctggaagcccgcccaaca	992
OY	841	TCTGCTGAGACCGTCTGGAAGCACTGATTACTTACATACAAATCTCAACTTCAGACC	900
Db	993	cttggtccaaagatctgcttagaanaacttgataaagtcatttagcagaatctcaactrcagagccc	1052
OY	901	GCCACACAAATTTAGAAACTGAGAGAGCTGCAGCAGAAAGTGTCTTACANAGGCGACCTTA	960
Db	1053	gtcaacaacaatlaaagaacttgagagagtlctgcacaaaagaatctctcaacaaggggcccaca	1112
OY	961	TCTGTCGACACCGGCCCATCTCTGAGAGAGAGGTGCTGTGAGCTGTTCGAAAACTTAATGA	1020
Db	1113	ctgtacagcagccggtcggaatgctcgtagagagagatctgtagagctgtgtcttcagaaacttaaga	1172
OY	1021	AGAGTGCCTTCTGTGTGTGAGCGCGACCGCTTCGATGCCATGCACCGGAGCCGGCCCTT AG	1080
Db	1173	aaagtgtcctcttggttgtagagcggtgagcccttgatgcgtccatctgacgtcggtccccctcg	1232
OY	1081	TCATCAAGACAGCGGTCTCACTTTTACACAGAAAGTACAGTGTGTGTCAAAATTTTCTAGAT	1140
Db	1233	tcatacaagaccgggtgtccagtlctacactcaactaaagtcaggtgtgtctgtcaagtcttcctgagt	1292
OY	1141	TGAATTTACAGCTTTAAATTTAAAGGTGTGATTTGATTAAGTCTCTGGAGATGTGCTGCC	1200
Db	1293	tgaattatcagcttaataatlaaagtgtgattatgcaaaagactcttggtggagctgtgtagctc	1352
OY	1201	TCAAGAGGCTCTCGGAATTTTAAATTTCTGCGCAGCAACCAAAAGTATGAACATGAGAG	1260
Db	1353	tcagagagatctccggaattttaaatactcttggtgacaaacaaagatgtagaacaatgtagag	1412

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2002, 14:00:57 ; Search time 1732.41 Seconds
(without alignments)
17795.810 Million cell updates/sec

Title: US-08-212-185-11

Perfect score: 2869
Sequence: 1 GCCGCGACCGCCGCGCGG.....AATTAAAAAAAAAAAAA 2869

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estlom:*
5: em_estlpl:*
6: em_estda:*
7: em_estro:*
8: em_estov:*
9: em_hlc:*
10: gb_estl:*
11: gb_estl2:*
12: gb_hlc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pin:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	766.6	26.7	947	11	B1156115 602903680
2	757	26.4	823	11	BG922983 602823860
3	752.2	26.2	788	11	BG862148 602795715
4	751.4	26.2	1036	11	B1156246 602903410
5	729.4	25.4	840	11	BG975502 602842871
6	718.8	25.1	731	11	BG920652 602826215
7	708.2	24.7	869	11	BG174177 602334170
8	702.6	24.5	889	11	BG175965 602337885
9	702.2	24.5	960	11	B1112360 602900049
10	700.2	24.4	717	11	BG976226 602846477
11	697	24.3	2634	12	AK018544 Mus muscu
12	694.4	24.2	719	11	B1108275 602902137

13	688.4	24.0	714	11	BG973911	BG973911 602843546
14	679	23.7	736	11	B1155944	B1155944 602904278
15	676.6	23.6	717	11	B1113023	B1113023 602898184
16	675.6	23.5	808	11	BG914848	BG914848 602813669
17	673	23.5	735	11	BG963048	BG963048 602828060
18	671.2	23.4	688	11	B1157660	B1157660 602920542
19	668	23.3	950	11	BG323650	BG323650 602421932
20	663.2	23.1	1181	11	BG174568	BG174568 602334411
21	657.6	22.9	744	11	B1219572	B1219572 602936656
22	652.6	22.7	841	11	BF141268	B1219572 602936656
23	652	22.7	666	11	B1157609	B1157609 602922581
24	651.8	22.7	1186	11	BE794248	BE794248 601591520
25	648.2	22.6	882	11	BF101658	BF101658 601753581
26	646	22.5	797	10	A1324586	A1324586 ms2304.y
27	643.8	22.4	1073	11	BE913188	BE913188 601668179
28	641.6	22.4	797	11	BG076355	BG076355 H3158G04-
29	627	21.9	1137	11	BF135796	BF135796 601782085
30	625.4	21.8	961	11	BF181575	BF181575 601808876
31	623.6	21.7	911	11	BF141093	BF141093 601786181
32	622	21.7	724	11	B1153867	B1153867 602870869
33	616.6	21.5	672	11	BG146440	BG146440 mab91f05-
34	614.4	21.4	616	11	BG864530	BG864530 602798560
35	613.8	21.4	724	11	BE794948	BE794948 601589709
36	613	21.4	736	10	AU132777	AU132777 AU132777
37	607.6	21.2	902	11	BG756243	BG756243 602713555
38	607.4	21.2	957	11	BF161870	BF161870 601766802
39	603.8	21.0	1123	11	BF143134	BF143134 601768806
40	602.6	21.0	676	10	BE374287	BE374287 601227646
41	599.4	20.9	802	11	BF162704	BF162704 601769359
42	597.2	20.8	772	11	BG914463	BG914463 602813202
43	594.6	20.7	931	11	BF142884	BF142884 601787952
44	591.6	20.6	656	10	BE284863	BE284863 601097521
45	587.8	20.5	782	11	BF584946	BF584946 602098830

ALIGNMENTS

RESULT 1
B1156115 947 bp mRNA
602903680P1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5033346 5',
LOCUS B1156115
ACCESSION B1156115
VERSION B1156115.1 GI:14616116
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1. http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM1092 row: P column: 19
High quality sequence stop: 772.
location/Qualifiers
1. 947
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:5033346"
/clone_lib="NIH_CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/lab_host="DH10B"

FEATURES

source

QY 1970 CCAGTCTGTAGAGCCATFACCAAGCAGACGCTGAACAACATGTCTATTGCTGAATCAT 2029
|||||
Db 241 CCAGTCTGTAGAGCCATFACCAAGCAGACGCTGAACAACATGTCTATTGCTGAATCAT 300
QY 2030 CATGGCCTTAAGATCATGATGCGCAACATCCGGTGTCTCCGCTGTACTCTTA 2089
|||||
Db 301 CATGGCCTTAAGATCATGATGCGCAACATCCGGTGTCTCCGCTGTACTCTTA 360
QY 2090 CCCCAGCATTTCCCAAGAGAGAGCATTTGGAAAGTACTGTAGCCCGAGAGCCAGAGCA 2149
|||||
Db 361 CCCCAGCATTTCCCAAGAGAGAGCATTTGGAAAGTACTGTAGCCCGAGAGCCAGAGCA 420
QY 2150 CCCCCAAGCCGAGCCAGTGTAGTGTCCCGCTACCTGAGAGCCAGTGTATCTGTGTAC 2209
|||||
Db 421 CCCCCAAGCCGAGCCAGTGTAGTGTCCCGCTACCTGAGAGCCAGTGTATCTGTGTAC 480
QY 2210 ACCAAGCAGCTGAGCAATACATTTGACCTGCCGATGTCCCGCAGCATTTAGATTCAAT 2269
|||||
Db 481 ACCAAGCAGCTGAGCAATACATTTGACCTGCCGATGTCCCGCAGCATTTAGATTCAAT 540
QY 2270 GATGCACTTTGCAATTAACGCGTGAAGTGTCTGAGCCCTCAGCAGAGAGCGCATTTGAGTC 2329
|||||
Db 541 GATGCACTTTGCAATTAACGCGTGAAGTGTCTGAGCCCTCAGCAGAGAGCGCATTTGAGTC 600
QY 2330 GCTCAGCTTTGACATGATGTGACCTCGAGTGTCTGACCTCCCGCATGTGAGAGCTGA 2389
|||||
Db 601 GCTCAGCTTTGACATGATGTGACCTCGAGTGTCTGACCTCCCGCATGTGAGAGCTGA 660
QY 2390 AACCCAGACCTGAGAGACGCTGACCTTGAGACACCTGCCCCCTCACCCTTAACGACG 2449
|||||
Db 661 AACCCAGACCTGAGAGACGCTGACCTTGAGACACCTGCCCCCTCACCCTTAACGACG 719
QY 2450 CGAACCCCATATGCTGTGAACCTCCTAACCTTGTGTGTCAGATTTTGTATTAATTC 2509
|||||
Db 720 -GAACCCATATGCTGTGAACCTCCTAACCTTGTGTGTCAGATTTTGTATTAATTC 777
QY 2510 CTACTTCTGCTATCTTTGGGCAATCTGGGCACTTTTAAAGA 2552
|||||
Db 778 CTACTTCTGCTATC-TTGGGCAATCTGGGCTTTTAAAGTGA 819
RESULT 3
Bg862148 788 bp mRNA EST 29-MAY-2001
LOCUS 602795715F1 NIH_CCAP_Mam4 Mus musculus cDNA clone IMAGE:4916931 5',
DEFINITION mRNA sequence.
ACCESSION Bg862148
VERSION Bg862148.1 GI:14212686
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 788)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9aps-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
plate: LHAM0837 row: n column: 04
High quality sequence start: 2
High quality sequence stop: 781.
Location/Qualifiers
1..788
/organism="Mus musculus"

/db_xref="taxon:10090"
/clone IMAGE:4916931
/clone_1db="NIH_CCAP_Mam4"
/tissue_type="tumor, gross tissue"
/lab_host="PH10B"
/note="Organ: mammary; Vector: pCMV-Sport6; Site_1: NCI;
Site_2: Sall; Cloned unidirectionally. Primer: Oligo dT.
Average insert 2.5 kb. Library constructed by Life
Technologies, catalog # 12018-016. Investigators providing
samples: Lothar Hennighausen/Priscilla Furth, NIH
Reference for transgenic model: Li et al., Cell Growth and
Differentiation 7, 3-11 (1996). Note: this is a NCI-CCAP
library."
BASE COUNT 209 a 191 c 216 g 172 t
ORIGIN

Query Match 26.2%; Score 752.2; DB 11; Length 788;
Best Local Similarity 98.7%; Pred. No. 3; 3e-162;
Matches 779; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

QY 1135 CTGAGTTGAATTATCAAGCTTAAATTAAGTGTGATGATAAAGACTCTGGGGATGTTG 1194
|||||
Db 1 CTGAGTTGAATTATCAAGCTTAAATTAAGTGTGATGATAAAGACTCTGGGGATGTTG 60
QY 1195 CTGCCCCTCAGAGGGTCTCGGAATTTTAACTTCTGGGCAAGACAAAAGTATGAGACA 1254
|||||
Db 61 CTGCCCCTCAGAGGGTCTCGGAATTTTAACTTCTGGGCAAGACAAAAGTATGAGACA 120
QY 1255 TGGAGAGCTTCAACAAGGCGACCTGCTGTGAGAGTTCAAGACCTGACTTATGAGGAGC 1314
|||||
Db 121 TGGAGAGCTTCAACAAGGCGACCTGCTGTGAGAGTTCAAGACCTGACTTATGAGGAGC 180
QY 1315 AGAGATGTGGGAATGAGAGCGCGTGGCAATGTGATGCTCTTGATGCTGACTGAGAGAC 1374
|||||
Db 181 AGAGATGTGGGAATGAGAGCGCGTGGCAATGTGATGCTCTTGATGCTGACTGAGAGAC 240
QY 1375 TGCACCTGATCACCCTTGAGAGCTGAGGTGTAGCACCAGGCTCAAGATTGACCTAGAGA 1434
|||||
Db 241 TGCACCTGATCACCCTTGAGAGCTGAGGTGTAGCACCAGGCTCAAGATTGACCTAGAGA 300
QY 1435 CCCAATCTTGGCAGTGTGTGATCTGCCAATCTGTGATGCCAATGCTTGGGCTAT 1494
|||||
Db 301 CCCAATCTTGGCAGTGTGTGATCTGCCAATCTGTGATGCCAATGCTTGGGCTAT 360
QY 1495 CAATCTCTAATACATCTGACCAATTAACCCCAAGAACTCTCTTCACTAAGC 1554
|||||
Db 361 CAATCTCTAATACATCTGACCAATTAACCCCAAGAACTCTCTTCACTAAGC 420
QY 1555 GGTACCTGGAGACCAATGCGCGAGGTGCTCAGCTTCTGCTCCACCA 1614
|||||
Db 421 GGTACCTGGAGACCAATGCGCGAGGTGCTCAGCTTCTGCTCCACCA 480
QY 1615 CCAAGCTGAGCATGAGAGCTGACCAAGCGTGGCTGTAAGCTCTTGAAGGCTG 1674
|||||
Db 481 CCAAGCTGAGCATGAGAGCTGACCAAGCGTGGCTGTAAGCTCTTGAAGGCTG 540
QY 1675 GTCATCTAGGCTGTACATGACATGGGCTAAATTCCTAAGAAACATGGGCTG 1734
|||||
Db 541 GTCATCTAGGCTGTACATGACATGGGCTAAATTCCTAAGAAACATGGGCTG 600
QY 1735 GCAAGAACTCTCTCTTCTGGGCTGTGCTAGACCAATATCATGACCTGTGAAAAGATATA 1794
|||||
Db 601 GCAAGAACTCTCTCTTCTGGGCTGTGCTAGACCAATATCATGACCTGTGAAAAGATATA 660
QY 1795 TCTTGGCCCTTTGGAAATGAAGGGTATCATGAGTGTTCATCAGCAAGAGCGGAGCGGG 1854
|||||
Db 661 TCTTGGCCCTTTGGAAATGAAGGGTATCATGAGTGTTCATCAGCAAGAGCGGAGCGGG 720
QY 1855 CCATCTTAAGCACAAAGCCCGGCGACCTTCTACTGCGCTTACAGCAGAGCAGCAAA 1913
|||||
Db 721 CCATCTTAAGCACAAAGCCCGGCGACCTTCTACTGCGCTTACAGCAGAGCAGCAAG 779

cdna library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.jnl.gov>

plate: LLAM10975 row: m column: 18

High quality sequence stop: 815.

FEATURES

Source

Location/Qualifiers

1..840

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone.lib="IMAGE:4978745"

/clone.lib="NIH_CGAP_Mam4"

/tissue.type="tumor, gross tissue"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert 2.5 kb. Library constructed by Life Technologies, catalog # 12018-016. Investigators providing samples: Lothar Hennighausen/Priscilla Furth, NIH

Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996). Note: this is a NCI-CGAP Library."

Library."

BASE COUNT

220 a 211 c 226 g 183 t

ORIGIN

Query Match 25.4% Score 729.4; DB 11; Length 840;

Best Local Similarity 95.98; Pred. No. 5.6e-157;

Matches 803; Conservative 0; Mismatches 26; Indels 8; Gaps 5;

QY 1802 CCTTGGATGAAGGATCATCATGCGTTTCATCAGCAGAGCGGAGCGCGCATCTCT 1861
 |||||||
 Db 2 CCGTTGGATGAAGGATCATCATGCGTTTCATCAGCAGAGCGGAGCGCGCATCTCT 61
 |||||||
 QY 1862 AAGCAAAAGCCCCGGGACCTTCTACTGCGCTTCAGCAGAGCAGCAAAAGAGAGG 1921
 |||||||
 Db 62 AAGCAAAAGCCCCGGGACCTTCTACTGCGCTTCAGCAGAGCAGCAAAAGAGAGG 121
 |||||||
 QY 1922 GGTCACTTTCTACTGCGTGAAGAAAGACATCAGTGGCAAGACCCAGATCCAGTCTGTGA 1981
 |||||||
 Db 122 GGTCACTTTCTACTGCGTGAAGAAAGACATCAGTGGCAAGACCCAGATCCAGTCTGTGA 181
 |||||||
 QY 1982 GCCATACCAAGCAGCAGCAGCTGAACAACATGTCTATTTGGTGAATCATCATGCGCTATAA 2041
 |||||||
 Db 182 GCCATACCAAGCAGCAGCAGCTGAACAACATGTCTATTTGGTGAATCATCATGCGCTATAA 241
 |||||||
 QY 2042 GATCATGGATGCGACCAACATCGTGTGTCCACTTGTCTACTCTTACCCCGCATTTCC 2101
 |||||||
 Db 242 GATCATGGATGCGACCAACATCGTGTGTCTCCACTTGTCTACTCTTACCCCGCATTTCC 301
 |||||||
 QY 2102 CAAGGAGAGGCACTTTGAAAGTACTGTAGGCCGAGAGCCAGAGCACCAGAGCCGA 2161
 |||||||
 Db 302 CAAGGAGAGGCACTTTGAAAGTACTGTAGGCCGAGAGCCAGAGCACCAGAGCCGA 361
 |||||||
 QY 2162 CCCAGGTAGTGTCTGCGCTGACCTGGAAGACCAATTCATCTGTGTGACACCAAGCAGCTG 2221
 |||||||
 Db 362 CCCAGGTAGTGTCTGCGCTGACCTGGAAGACCAATTCATCTGTGTGACACCAAGCAGCTG 421
 |||||||
 QY 2222 CAGCAATACCTTACCTGCGATGTCGCCCGCGCATTTAGATTCATTCATGATGAGTTGG 2281
 |||||||
 Db 422 CAGCAATACCTTACCTGCGATGTCGCCCGCGCATTTAGATTCATTCATGATGAGTTGG 481
 |||||||
 QY 2282 AAATAACGGTGAAGGTGTGAGCCCTCAGCAGAGAGGCGAGTTTGTAGTGTGACGTTTGA 2341
 |||||||
 Db 482 AAATAACGGTGAAGGTGTGAGCCCTCAGCAGAGAGGCGAGTTTGTAGTGTGACGTTTGA 541
 |||||||
 QY 2342 CATGATCTGACCTCGGAGTGTGTACTCTCCCATGATGATGATGATGATGATGATGATGAT 2401
 |||||||
 Db 542 CATGATCTGACCTCGGAGTGTGTACTCTCCCATGATGATGATGATGATGATGATGATGAT 601
 |||||||
 QY 2402 CAGAGAGTGTGAGACACCTGTGCCCC -GTGCTTCAACCCCTTAAGCAGCCAGCCCATTA 2460
 |||||||

Db 602 CAGAGACCTTCACTTGAAGACACCTGCCCGGTTGCTCCAAACCTTAAGCAGGCGAA-CCCATTA 660

QY 2461 TCGTCTAAACCTCTTAACCTTTGGTGTCCAGATT--TTTTTTTTTAATTTCTACTTCTG 2518
 |||||||

Db 661 TCGTCTAAACCTCTTAACCTTTGGTGTCCAGATTATTTGTTAAATTTCTACTTCTG 720
 |||||||

QY 2519 CTATCTTGGGCATCTGGGCACCTTTTAAAGAGAGAATGATGATGATGATGATGATGATGAT 2578
 |||||||

Db 721 GTATCTTTGGGCATCTGGGCACCTTTTAAATGAGAGAATGATGATGATGATGATGATGAT 780
 |||||||

QY 2579 ACTGTATGTAAGAGAGACACCTCTGAGTGTGGATGGGCTGAGACGAGAAG 2635
 |||||||

Db 781 --CTTATGTAAAGAGAGAGAC--TTGAGTCTGGGATGGGATGAGAGCGAAAGGG 833
 |||||||

RESULT 6

BG920652 731 bp mRNA EST 05-JUN-2001

LOCUS 602826215F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4954946 5',

DEFINITION mRNA sequence.

ACCESSION BG920652

VERSION BG920652.1 GI:14301128

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>, National Institutes of Health, Mammalian Gene Collection (MGC)

AUTHORS Unpublished (1999)

TITLE Contact: Robert Strausberg, Ph.D.

JOURNAL Email: cgabs-remail.nih.gov

COMMENT Tissue Procurement: Jeffrey Green M.D.

cdna library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.jnl.gov>

plate: LLAM10916 row: n column: 03

High quality sequence start: 2

High quality sequence stop: 725.

Location/Qualifiers

1..731

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone.lib="IMAGE:4954946"

/clone.lib="NCI_CGAP_Mam6"

/sex="female, virgin"

/tissue.type="inflicting ductal carcinoma"

/dev_stage="5 months"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-Sport6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 191 a 183 c 198 g 158 t 1 others

ORIGIN

Query Match 25.1% Score 718.8; DB 11; Length 731;

Best Local Similarity 98.9% Pred. No. 1.6e-154;

Matches 723; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 899 CCGCAACAAATTAAGAACTGAGAGAGCTGCAGCAGAAAGTGTCTACAAAGGCGACCC 958
 |||||||

Db 1 CCGCAACAAATTAAGAACTGAGAGAGCTGCAGCAGAAAGTGTCTACAAAGGCGACCC 60
 |||||||

QY 959 TATGTGTG TGCAGCC TCCCATGCTGAGAGAGAGATGTGTGAGCTGTTGAGAACTTAAT 1018
 |||||||

Db 61 TATGTGTGAGAGAGAGAGAGAGAGATGTGTGAGCTGTTGAGAACTTAAT 120
 |||||||

```
OY 1019 GAAGAGTGCCTTCGTGGTGGAGCGGACGCCCTGCATGCCATGACCCGAGCGGCCCTT 1078
|||||
DB 121 GAAGAGTGCCTTCGTGGTGGAGCGGACGCCCTGCATGCCATGACCCGAGCGGCCCTT 180
|||||
OY 1079 AGTATCAAGCTGTGTGCCAGTTTACCAGAAAGTCAAGTGTGTGTCAATTTCTGA 1138
|||||
DB 181 AGTATCAAGCTGTGTGCCAGTTTACCAGAAAGTCAAGTGTGTGTCAATTTCTGA 240
|||||
OY 1139 GTTGAATTTAGCTTAATAATTAAGTGTGTGTGTGAATGAAGTCTGGGATGTTGTGC 1198
|||||
DB 241 GTTGAATTTAGCTTAATAATTAAGTGTGTGTGAATGAAGTCTGGGATGTTGTGC 300
|||||
OY 1199 CCTCAGAGGCTGTGGAAATTTAACTTTCTGGGACGACCAACAAAGATGATGATGGA 1258
|||||
DB 301 CCTCAGAGGCTGTGGAAATTTAACTTTCTGGGACGACCAACAAAGATGATGATGGA 360
|||||
OY 1259 GGAGCTTAACAGGCGACGCTGTCTGAGAGTTCACAGCCTGACCTTAGGAGAGAG 1318
|||||
DB 361 GGAGCTTAACAGGCGACGCTGTCTGAGAGTTCACAGCCTGACCTTAGGAGAGAG 420
|||||
OY 1319 ATGTGGGAATGAGAGCGCTGCCAATTTGTGATGCTCTTGTGATCGTGAAGAGCTGCA 1378
|||||
DB 421 ATGTGGGAATGAGAGCGCTGCCAATTTGTGATGCTCTTGTGATCGTGAAGAGCTGCA 480
|||||
OY 1379 CCTGATCACTTCGAGAGTGTGATGTACACCAAGGCTCAAGATTTGACAGAGACCA 1438
|||||
DB 481 CCTGATCACTTCGAGAGTGTGATGTACACCAAGGCTCAAGATTTGACAGAGACCA 540
|||||
OY 1439 CTCCTGCGCAGTGTGTGTGTGTCTCCAGATCTGTGATGCTGATGCTGAGAGCTCAAT 1498
|||||
DB 541 CTCCTGCGCAGTGTGTGTGTGTCTCCAGATCTGTGATGCTGATGCTGAGAGCTCAAT 600
|||||
OY 1499 CCTGTGATTAACATGCTGACCAATAACCCCAAGAGCTGAACCTTCTCACTAAAGCCGC 1558
|||||
DB 601 CCTGTGATTAACATGCTGACCAATAACCCCAAGAGCTGAACCTTCTCACTAAAGCCGC 660
|||||
OY 1559 AATTGGAACCTGTGGACCAAGTGGCCGAGGTGCTCAGCTGGCAGTTCTGTCCACCAACA 1618
|||||
DB 661 AATTGGAACCTGTGGACCAAGTGGCCGAGGTGCTCAGCTGGCAGTTCTGTCCACCAACA 720
|||||
OY 1619 GCGAGGCGCTGA 1629
|||||
DB 721 GCGAGGCGCTGA 731
|||||
```

```
FEATURES
source
Location/Qualifiers
1..869
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4457248"
/tissue_type="NCI CGAP Maml"
/tissue_type="tumor, biopsy sample"
/dev_stage="3 months, virgin"
/lab_host="IMH08"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT 245 a 220 c 253 g 151 t
ORIGIN
Query Match 24.7% Score 708.2; DB 11; Length 869;
Best Local Similarity 97.1% Pred. No. 4.1e-152;
Matches 764; Conservative 0; Mismatches 18; Indels 5; Caps 4;
OY 407 TCGCTCTCCAGAGAGCGAGCGGACGCCAGCCAGAGGGGCCAGGCCAGCCACCCAC 466
|||||
DB 1 TCGCTCTCCAGAGAGCGAGCGGACGCCAGCCAGAGGGGCCAGGCCAGCCACCCAC 60
|||||
OY 467 AGCGCGCTAGTGAGAGAGAGAGAGAGATGTTGAGAGCATCTTCAGATGTCGGAA 526
|||||
DB 61 AGCGCGCTAGTGAGAGAGAGAGAGAGATGTTGAGAGCATCTTCAGATGTCGGAA 120
|||||
OY 527 GCGAGTGCAGATCTAGAAACGAAATGAAGTGTGTGAGAACTCCAGAGAGCTTTGA 586
|||||
DB 121 GCGAGTGCAGATCTAGAAACGAAATGAAGTGTGTGAGAACTCCAGAGAGCTTTGA 180
|||||
OY 587 TTTCACTACAAACCCCTCAAGAGAGAGAGAGATGAGATGATGTAATGGAACAACA 646
|||||
DB 181 TTTCACTACAAACCCCTCAAGAGAGAGAGAGATGAGATGATGTAATGGAACAACA 240
|||||
OY 647 GTCTGTGACAGACAGAGAGATGACAGCTGTGAACAGATGCTCACAGCCCTGACAGAT 706
|||||
DB 241 GTCTGTGACAGACAGAGAGATGACAGCTGTGAACAGATGCTCACAGCCCTGACAGAT 300
|||||
OY 707 GCGGAGAGCATGTTGAGTGTGAGCTGTGGGGGCTCTTGTGACGAATGAGTACGTGAGAA 766
|||||
DB 301 GCGGAGAGCATGTTGAGTGTGAGCTGTGGGGGCTCTTGTGACGAATGAGTACGTGAGAA 360
|||||
OY 767 GACACTGACTGTGAAGAGCTGTGACTGTGAAGAGCGGCCAGAGATCGCTGTGATCGG 826
|||||
DB 361 GACACTGACTGTGAAGAGCTGTGACTGTGAAGAGCGGCCAGAGATCGCTGTGATCGG 420
|||||
OY 827 AGGCCCTCCCAACATCTGCTGTGAGACCGCTGTGAAACCTGGAATCTTATTAAGCAATC 886
|||||
DB 421 AGGCCCTCCCAACATCTGCTGTGAGACCGCTGTGAAACCTGGAATCTTATTAAGCAATC 480
|||||
OY 887 TCAACTTGAGACCCGCCCAACAAATTAAGAACTGAGAGAGTGTGAGCAAGAAAGTGTCTTA 946
|||||
DB 481 TCAACTTGAGACCCGCCCAACAAATTAAGAACTGAGAGAGTGTGAGCAAGAAAGTGTCTTA 540
|||||
OY 947 CAAGGCGACCTATCTGTGACAGACACCGCCCATCTGTGAGAGAGAGATCTGTGAGACTTT 1006
|||||
DB 541 CAAGGCGACCTATCTGTGACAGACACCGCCCATCTGTGAGAGAGAGATCTGTGAGACTTT 600
|||||
OY 1007 CAGAAACTTAATGAAGAGTGTCTGTGTGAGAGGGGAGGCCCTGCATGCCATGACACCC 1066
|||||
DB 601 CAGAAACTTAATGAAGAGTGTCTGTGTGAGAGGGGAGGCCCTGCATGCCATGACACCC 660
|||||
OY 1067 GGAGCGGCCCTTAGTCATCAAGACT-GGTGTCCAGTTTACC-ACGAAAGTCAAGTTGCTGTG 1124
|||||
DB 661 GGAGCGGCCCTTAGTCATCAAGACTGTGTGTCCAGTTTACCACGAAAGTCAAGTTGCTGTG 720
|||||
OY 1125 GTCAATTTCTGCA-GTTGCAATTATCAGCTT--AAATTTAAAGTGTGATGTAATAAC 1181
|||||
DB 721 GTCAATTTCTGAGGTGGAATTAATCAGGTTTAAATTAACAGGGGTGCTATGATCAAGAC 780
|||||
```

QY 1182 TCTGGGG 1188
|||||
Db 781 TCTGGGG 787

RESULT 8
Bg175965 889 bp mRNA EST 06-FEB-2001
LOCUS 602337885F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4461020 5',
DEFINITION mRNA sequence.
ACCESSION Bg175965
VERSION Bg175965.1 GI:12682668
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Mammalia: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Eukaryota: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.

REFERENCE 1 (bases 1 to 889)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Inceye Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
plate: LLAM10263 row: i column: 21
High quality sequence stop: 714.
Location/Qualifiers
1..889
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4461020"
/clone_1lb="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="3 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 221 a 232 c 242 g 194 t
ORIGIN

Query Match 24.5%; Score 702.6; DB 11; Length 889;
Best Local Similarity 96.9%; Pred. No. 7.8e-151;
Matches 770; Conservative 0; Mismatches 19; Indels 6; Gaps 5;

QY 1177 GCAAGAAACATGCTGCGCAAGGCTTCTCTTGTGCTGTAGACATATCATCG 1776
|||||
Db 1 GCAAGAAACATGCTGCGCAAGGCTTCTCTTGTGCTGTAGACATATCATCG 60

QY 1777 ACCTTGAAAAAGTATATCTTGGCCCTTGGAAATGAAGGTACATATGGCTTCATCA 1836
|||||
Db 61 ACCTTGAAAAAGTATATCTTGGCCCTTGGAAATGAAGGTACATATGGCTTCATCA 120

QY 1837 GCAAGAGCGGAGCGGCGCATCTTAAGCACAAAGCCCCGGGCACTTCTTACTGGCCT 1896
|||||
Db 121 GCAAGAGCGGAGCGGCGCATCTTAAGCACAAAGCCCCGGGCACTTCTTACTGGCCT 180

QY 1897 TCACGAGACAGACGAAGAAGAGGGGTCACTTTCACCTTGGTGGAAAAAGACATCAAGT 1956
|||||
Db 181 TCACGAGACAGACGAAGAAGAGGGGTCACTTTCACCTTGGTGGAAAAAGACATCAAGT 240

QY 1957 GCAAGACCCAGATCAGTCTGTAGAGCATACACCAAGCAGCAGCTGAACAATGTCTAT 2016
|||||

Db 241 GCAAGACCCAGATCCAGTCTGTAGAGCCATACACCAAGCAGCAGCTGACACATGTCTAT 300

QY 2017 TTGCTGAATATCATCATGCGCTATATAGATGCATGCGACCAACATCCTGGTCTCCAC 2076
|||||
Db 301 TTGCTGAATATCATCATGCGCTATATAGATGCATGCGACCAACATCCTGGTCTCCAC 360

QY 2077 TTGCTACCTCTACCCCGACATTCGCCAAGAGAGGAGGATTTGGAATACGTAGAGGCCCG 2136
|||||
Db 361 TTGCTACCTCTACCCCGACATTCGCCAAGAGAGGAGGATTTGGAATACGTAGAGGCCCG 420

QY 2137 AGAGCCGAGAGCACCAGCCGAGAGCCGAGGTAAGTGTGCCCCCTGACCTGAAGACCAAGT 2196
|||||
Db 421 AGAGCCGAGAGCACCAGCCGAGAGCCGAGGTAAGTGTGCCCCCTGACCTGAAGACCAAGT 480

QY 2197 TCATCTCTGTGACACCAAGCAAGCACTGACGAATATACCTTGCATGCTCCCGCCGCA 2256
|||||
Db 481 TCATCTCTGTGACACCAAGCAAGCACTGACGAATATACCTTGCATGCTCCCGCCGCA 540

QY 2257 CTTAGATTCATTCATGATGACGATTTGGAATTAACGGGTGAAGGCTGAGCCCTGACGAGAG 2316
|||||
Db 541 CTTAGATTCATTCATGATGACGATTTGGAATTAACGGGTGAAGGCTGAGCCCTGACGAGAG 600

QY 2317 GCGAGTTTGAATGCTGCTGACGCTTTTGACATGATC-TGACCTGAGAGT-6CTACCTCCC 2374
|||||
Db 601 GCGAGTTTGAATGCTGCTGACGCTTTTGACATGATC-TGACCTGAGAGT-6CTACCTCCC 660

QY 2375 CATGTGAGAGCTTAACCAAGCAAGCTGCAGAGAGCTGACTTGAACACCTTCCCGCT 2434
|||||
Db 661 CATGTGAGAGCTTAACCAAGCAAGCTGCAGAGAGCTGACTTGAACACCTTCCCGCT 718

QY 2435 CCACCCCTTAAGCAGCCGACACCCATATGCTGTAAGCTTCTTAAGCTTGTGCTGACATTT 2494
|||||
Db 719 CCACCCCTTAAGCAGCCGACACCCATATGCTGTAAGCTTCTTAAGCTTGTGCTGACATTT 776

QY 2495 TTTTCTTTTAAATTC 2509
|||||
Db 777 TTTAATTCCTACTTC 791

RESULT 9
B1112360 960 bp mRNA EST 26-JUN-2001
LOCUS B1112360
DEFINITION 602900049F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5029548 5',
LOCUS B1112360
DEFINITION mRNA sequence.
ACCESSION B1112360
VERSION B1112360.1 GI:14563261
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Mammalia: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Eukaryota: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.

REFERENCE 1 (bases 1 to 960)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Inceye Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
plate: LLAM11083 row: b column: 13
High quality sequence stop: 789.
Location/Qualifiers
1..960
/organism="Mus musculus"
/strain="C57/B6"
/db_xref="taxon:10090"
/clone="IMAGE:5029548"
/clone_1lb="NCI_CGAP_Mam5"

/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: PCMV-SPORT6; Site:1: Sal1;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lochar Hennighausen/Kodin Humphreys,
NIH"

BASE COUNT 247 a 258 c 260 g 195 t
ORIGIN

Query Match 24.5%: Score 702.2: DB 11: Length 960:
Best Local Similarity 94.3%: Pred. No. 9.5e-151:
Matches 784: Conservative 0: Mismatches 38: Indels 9: Gaps 5:

OY 1436 CCACCTCTGCGCAGTGTGTGATCTCAACATCTGTACATGCCAAATGCTTGGGCATC 1495
|||||
DB 4 CCACCTCTGCGCAGTGTGTGATCTCAACATCTGTACATGCCAAATGCTTGGGCATC 63
OY 1496 AATCTGTGTATACATGCTGACCAATAACCCCAAGACGTGAATCTTCACTAAGCC 1555
|||||
DB 64 AATCTGTGTATACATGCTGACCAATAACCCCAAGACGTGAATCTTCACTAAGCC 123
OY 1556 GCCAATTTGAACTGGGACCAAGTGGCCGAGTGTGCTCAGCTGGCAATTTCTGTCACACCAC 1615
|||||
DB 124 GCCAATTTGAACTGGGACCAAGTGGCCGAGTGTGCTCAGCTGGCAATTTCTGTCACACCAC 183
OY 1616 CAAGCGAGGCTGAGCATCGAGCAGCTGACACGCTGGCTGAGAAGCTCTAGGGCCTGG 1675
|||||
DB 184 CAAGCGAGGCTGAGCATCGAGCAGCTGACACGCTGGCTGAGAAGCTCTAGGGCCTGG 243
OY 1676 TGTGAACTACTCAGGCTGTAGATCAGATGGCTAAATTTCTGCAAGAAAACATGGCTGG 1735
|||||
DB 244 TGTGAACTACTCAGGCTGTAGATCAGATGGCTAAATTTCTGCAAGAAAACATGGCTGG 303
OY 1736 CAAGGCGTCTCTGCTGGGCTGGCTAGACAAATATCATGACCTTGTGAAAAAGTATAT 1795
|||||
DB 304 CAAGGCGTCTCTGCTGGGCTGGCTAGACAAATATCATGACCTTGTGAAAAAGTATAT 363
OY 1796 CTGGGCGCTTGGAAATGAAGGGATACATCATGGCTTTCATCAGCAAGAGCGGGAGCGGC 1855
|||||
DB 364 CTGGGCGCTTGGAAATGAAGGGATACATCATGGCTTTCATCAGCAAGAGCGGGAGCGGC 423
OY 1856 CATCTTAAGACAAAGCCCGGCGACCTTCTACTGCGCTTTCAGCGAGACGACCAAGA 1915
|||||
DB 424 CATCTTAAGACAAAGCCCGGCGACCTTCTACTGCGCTTTCAGCGAGACGACCAAGA 483
OY 1916 AGGAGGGTCACTTTCACCTTGGGTGGAAGAGCATAGTGGCAAGACCCAGATCCAGTC 1975
|||||
DB 484 AGGAGGGTCACTTTCACCTTGGGTGGAAGAGCATAGTGGCAAGACCCAGATCCAGTC 543
OY 1976 TGTAGACCATATACCAAGCAGCAGTGAACACATCTCATTTGCTGTAATCATATGGG 2035
|||||
DB 544 TGTAGACCATATACCAAGCAGCAGTGAACACATCTCATTTGCTGTAATCATATGGG 603
OY 2036 CTATAAATCATGATGATCGACCAACATCGTGTGTCT-CCACTGTCTACTCTTACCCCG 2094
|||||
DB 604 CTATAAATCATGATGATCGACCAACATCGTGTGTCTCCCACTGTCTTACTCTTACCCCG 663
OY 2095 ACATTTCCCAAGAGAGGAGCATTTT--GGAAAGTACTGTA-GGCCGAGACGCCAGAGCAACC 2151
|||||
DB 664 ACATTTCCCAAGAGAGGAGCATTTGGAAGAGTCTCTGTAAGGGCCGGAAGACGAGACGACC 723
OY 2152 CCGAAGCCGAGCCAGAGTGTGTCGCCCGTACTGGAAGACCAA--GTTCAATCTGTGTGAC 2209
|||||
DB 724 GCGAAGCCGAGCCAGTGTGTCGCCCGTACTGGAAGAGTCTCTGTAAGGGCCGGAAGACGAGAC 783
OY 2210 ACCAAGACCC--TGCAGCAATACGATTAAGCTGGCGATGTCGCCCGGAC 2257
|||||
DB 784 ACCAAGACCCGTGAGGCAATATACCATTTGACCTGCGATGTGTCTCCCGCCG 834

RESULT 10
BG976226 717 bp mRNA EST 12-JUN-2001
LOCUS 602846477F1 NC1_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4977115 5',
DEFINITION mRNA sequence.
ACCESSION BG976226
VERSION BG976226.1 GI:14363863
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
1 (bases 1 to 717)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-f@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: LLAM10971 row: 1 column: 20
High quality sequence stop: 699.

FEATURES
source
1..717
Location/Qualifiers
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_1fb="NC1_CGAP_Mam6"
/sex="Female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: PCMV-SPORT6; Site:1: Sal1;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 197 a 175 c 187 g 158 t
ORIGIN

Query Match 24.4%: Score 700.2: DB 11: Length 717:
Best Local Similarity 99.4%: Pred. No. 2.9e-150:
Matches 713: Conservative 0: Mismatches 3: Indels 1: Gaps 1:

OY 857 GGAAGCTGATTAATCTTATAGAGAAATCTCAACTTCAGACCGGCAACAAATTAAGAA 916
|||||
DB 1 GGAAGCTGATTAATCTTATAGAGAAATCTCAACTTCAGACCGGCAACAAATTAAGAA 60
OY 917 ACTGAGAGAGCTGACGAGGAAGTGTCTCTACAGGCGACCTTATCTCATCAAGAGTGTGT 976
|||||
DB 61 ACTGAGAGAGCTGACGAGGAAGTGTCTCTACAGGCGACCTTATCTCATCAAGAGTGTGT 120
OY 977 CATGCTGAGAGAGAGATCGTGAAGCTGTTTCAAGAACTTAATGAAGAGTGTCTGTGCT 1036
|||||
DB 121 CATGCTGAGAGAGAGATCGTGAAGCTGTTTCAAGAACTTAATGAAGAGTGTCTGTGCT 180
OY 1037 GAGACGGGACCCCTGCAATGCCCATGACACCGGACCGGCGCTTATCTCATCAAGAGTGTGT 1096
|||||
DB 181 GAGACGGGACCCCTGCAATGCCCATGACACCGGACCGGCGCTTATCTCATCAAGAGTGTGT 240
OY 1097 CCAGTTTACCAAGCAAGTCAAGTGTGTCGAATTTCTGAGTGTGAATATACAGCTTAA 1156
|||||
DB 241 CCAGTTTACCAAGCAAGTCAAGTGTGTCGAATTTCTGAGTGTGAATATACAGCTTAA 300
OY 1157 AATTAAAGTGTGATTAAGAACTCTGGGAGTGTGCTGCTCCTCAGAGGGTCTCGGAA 1216
|||||
DB 301 AATTAAAGTGTGATTAAGAACTCTGGGAGTGTGCTGCTCCTCAGAGGGTCTCGGAA 360

QY 1217 ATTTACATTCTGGGACGAACAAAGTCATGACATGAGAGCTCTAACACGCGAG 1276
|||||
Db 361 ATTAAACATTCTGGGACGAACAAAGTCATGACATGAGAGCTCTAACACGCGAG 420
QY 1277 CCTGTGCGAGAGTTCAGAGCCCTGAGCCCTGAGGAGCAGAGATGCTGGAAATGGAGCGG 1336
|||||
Db 421 CCTGTGCGAGAGTTCAGAGCCCTGAGCCCTGAGGAGCAGAGATGCTGGAAATGGAGCGG 480
QY 1337 TGCCCAATTCGATCCCTCTGATGCTGAGCTGAGAGAGCTGACCTGATCCTTGGAGAC 1396
|||||
Db 481 TGCCCAATTCGATCCCTCTGATGCTGAGCTGAGAGAGCTGACCTGATCCTTGGAGAC 540
QY 1397 TGAGGTGACCAAGAGCGCTCAAGATTGACCTGAGAGACCCACTCTTGGCAATTTGCT 1456
|||||
Db 541 TGAGGTGACCAAGAGCGCTCAAGATTGACCTGAGAGACCCACTCTTGGCAATTTGCT 600
QY 1457 GATGTCCACATCTGTGATGATGAGATGCTGGGATCAAA-TGCTGTGATTAACATG 1515
|||||
Db 601 GATGTCCACATCTGTGATGATGAGATGCTGGGATCAAA-TGCTGTGATTAACATG 660
QY 1516 TGACCAATTAACCCAGAACGTAAGCTTCTTCACTAAGCCGCAATTTGAACCTGGG 1572
|||||
Db 661 TGACCAATTAACCCAGAACGTAAGCTTCTTCACTAAGCCGCAATTTGAACCTGGG 717
RESULT 11
AK018544 2634 bp mRNA HMC 05-JUL-2001
LOCUS Mus musculus adult male colon cDNA, RIKEN full-length enriched
DEFINITION library, clone:9030611B09, full insert sequence.
ACCESSION AK018544
VERSION AK018544.1 GI:12858297
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male colon cDNA to mRNA,
clone:9030611B09.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Methods in enzymology. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome research. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE
AUTHORS 3 (bases 1 to 2634)
Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuami,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,
Yamamoto,K., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujimake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multiplexillary sequencer
JOURNAL Genome research. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE
AUTHORS 4 (bases 1 to 2634)
The Riken Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5 (bases 1 to 2634)

AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,
Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F.,
Imotani,K., Ishii,Y., Itoh,M., Iwama,M., Kato,H., Kawai,J.,
Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T.,
Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y.,
Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H.,
Sasaki,D., Shibata,K., Shibata,Y., Shihagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A.,
Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gs.c.riken.go.jp,
URL:ftp://genome-gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT Please visit our web site (<http://genome-gsc.riken.go.jp/>) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
15' GAGAGAGAGAGATTCGACGATTAATTAATGATCCGCCCCCCC 3'. cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 10.0 and subtraction to
Rot = 185.2. Second strand cDNA was prepared with the primer
adapter of sequence 15'
GAGAGAGAGATTCGACGATTAATTAATGATCCGCCCCCCC 3'. cDNA was cleaved
with BamHI and XhoI. Vector: a modified pBluescript KS(+) after
bulk excision from Lambda F1C 1. Cloning sites, 5' end: SalI; 3'
end: BamHI. Host: DH10B.
FEATURES
Source
Location/Qualifiers
1..2634
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="MGI:103063"
/db_xref="MGI:1907734"
/clone="9030611B09"
/sex="male"
/tissue_type="colon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
302..2551
/note="putative"
/codon_start=1
/protein_id="BAB31265.1"
/db_xref="GI:12858298"
/translation="MSQMFLEQDLSKFLRQVHLYDPSFMEIRQYLAQMLEKQWE
HAADVSEFATREHDLISOLDQYSRSLLENFLQINIRKSKRNLDNOEDPVQMS
MIYNCLKEERKILLENQKROEGNIONTVMLDKQKELSKRNLDNOEDPVQMS
KTLLELDEYDFKCTSONREGANGVAKSDQKQELLLHMFIMLNKKEIILHKR
ELNLSILDTONTLLINDELVEKRRQSGACIGPNACLDLQSFITVAETLQDIIRO
LKTLELEOKFTPEPDIYTKNKLVSDFELFQOLIOSSFVVRQCMPTHPORPLV
LKTQVOTVTKRLTLKLOELNYNLKAVSFDKQVNEKNTYKGFKNIICTHNVVM
EESTNGSLAEFRILQIKOKNGNRTNEPRLIYTERLSLSPETQICQGLVYDET
TSLPVPVYISNVSQPSMASILNMYLTYTERNLISFLNPLCAWMSQSLSEVLSQFS
VTKRGLNADQSLNIGERCLDGNAPDGLIPTWTRCKNINDKNSFWPWIDTILELK
KHLCLVANDQCIWGFISKERERALLKQDQPTFLRPSSESRGALITFWVERSONG
EPDFHAYNDKELSAVTPPDILIRNRYVMAEENIPNPLKTYLIPNIDKHALGKYS
RPEKAPPEMELDPRKGYIKTELISVSEVHPSRLQTTDNLPLPSPPEPMESRIYVP
EFDMSMTV"

BASE COUNT 729 a 658 c 658 g 589 t
ORIGIN
Query Match 24.3% Score 697; DB 12; Length 2634;
Best Local Similarity 60.4%; Pred. No. 1.2e-149;

Matches 1248: Conservative 0: Mismatches 795: Indels 24: Gaps 5:				
OY	66	AGATGGCTCTACTGGAAACAGCTGCAGACAGCTGGACACAGCTACTGTGAAGCAGCTGCAC	125	
Db	299	AGATGTCTACAGTGGTTCAGCTTCCAGCTTGACCTGACCTGACCTGAGCTGAGCTGAC	358	
OY	126	CAGCTTACAGGACAGCTTCCCATGAGCTCCGGCAGCTTCTGGCACCCTTGATGAG	185	
Db	359	CAGCTTACAGGACAGCTTCCCATGAGCTTCCCATGAGCTTCCCATGAGCTTCCCATGAG	418	
OY	186	AGCAGACAGCTGGGATATGAGCCAGCAGCAAGAGTACATGACATGCTGATGATTAAT	245	
Db	419	AGCAGACAGCTGGGATATGAGCCAGCAGCAAGAGTACATGATGATGATGATGATGATGAT	478	
OY	246	CTCTTGGGTAATTTGACCAAGCAATATAGCCGATCTCTGCAAGAGTCCATGCTCTAT	305	
Db	479	CTCTCTCTCAGAGCTGAGCAGCAGTACAGCCTTTCTCTGAGAAATTTCTTCTTCTT	538	
OY	306	CAGCAGACCTTGGAGAAATCAAGCAGATTTCTGAGAGCAGTATCTTGAAGCCAAATG	365	
Db	539	CAGCAGACCTTGGAGAAATCAAGCAGATTTCTGAGAGCAGTATCTTGAAGCCAAATG	598	
OY	366	GAATTTGCCCGATCGTGCCCGATGCTGTGGAGAGTCTGCTCTCCAGACGGCA	425	
Db	599	CAGATGTCATGATCTACACTGCTCTGAAGGAGAAAGAAATTTGAAATATGCC	658	
OY	426	GCCAGGCGACCCAGCAGAGGGGGCCAGCCACCCACCCAGCCGCGCTAGTGCAGAG	485	
Db	659	CAAGATTTTAAATCA-----GGCCAGAGGAGAAATTTTCAAGAACCTGTGATTTGAT	712	
OY	486	AAGCAGAGATGTTGGAGCAGCATCTTCAAGATGTCGGAAGCAGTGCAGATCTAGAA	545	
Db	713	AAGCAGAGATGTTGGAGCAGCATCTTCAAGATGTCGGAAGCAGTGCAGATCTAGAG	772	
OY	546	CAGAAATGAAGTGTGTGAGAACCTCCAGAGCAGCTTTGATTTTCAACTACAAACCTTC	605	
Db	773	CAGAAATGAAGTGTGTGAGAACCTCCAGAGCAGCTTTGATTTTCAACTACAAACCTTC	832	
OY	606	AAGAGCCAGAGACATGCAAGATCTCAATGGAACCAACAGCTGTGAGCAGCAGAG	665	
Db	833	CAGAGCAGAGAG-----GTGAAGCCAAATGCTGTGCGAAGCAGCCCAAAACAGGAA	886	
OY	666	ATGAGAGAGCTGGAACAGATGCTCAGAGCCCTGAGCAGATGCGAGAGAGCATTTGTAGT	725	
Db	887	CAGCTGCTGCTCCAGAGATTTTATGCTGTGACATTAAGAGAAAGAGATTAATTCAC	946	
OY	726	GAGCTGCGGGCTCTTGTGCAATGAGTACGTGCAAGAGACACTGACTGATGAGAG	785	
Db	947	AAATTCAGTGAGTTGCTGAATTTCCATGAGCTCACTCAGAACACTGTGATTAATGACAG	1006	
OY	786	CTGCTGCTGAGAGAGAGGAGAGATGCTGATGAGAGAGAGAGAGAGAGAGAGAGAG	845	
Db	1007	CTGCTGCTGAGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1066	
OY	846	CTGAGCCTCTGGAAGAACTGATTAATTCATTAAGCAGAAATCTCAATTCAGACCCGCCAA	905	
Db	1067	CTGATGATGAG	1126	
OY	906	CAAAATTAAGAACTGAGAGAGCTGAGCAAGAGTGTCTTACAGAGGAGAGAGAGAG	965	
Db	1127	CAGCTTAAAGAGCTGAG	1186	
OY	966	CAGAGCCGCCCTGAG	1025	
Db	1187	AAAAACAAG	1246	
OY	1026	GCTTCTGCTGAG	1085	
Db	1247	TCTTCTGCTGAG	1306	
OY	1086	AAGAGCTGCTCAGATTAACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1145	
Db	1307	AAGAGCTGCTCAGATTAACAGCTGCTCAAGCTGAGAGAGAGAGAGAGAGAGAGAG	1366	

OY	1146	TATCAGCTTAAATTAAGTGTGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1205	
Db	1367	TATCAGCTTAAATTAAGTGTGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1426	
OY	1206	GAGTCTCGAAATTTAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1265	
Db	1427	GAGTCTCGAAATTTAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1486	
OY	1266	AACAAGCCGAG	1325	
Db	1487	ACCAAG	1540	
OY	1326	AATGAG	1385	
Db	1541	AAGCCTGGAAG	1600	
OY	1386	ACCTTGCAG	1445	
Db	1601	AGCTTGAAG	1660	
OY	1446	CCAGTGTGATCTGCAATCTGTCAGATGTCAGAGAGAGAGAGAGAGAGAGAGAG	1505	
Db	1661	CTGTGCTGATCTGCAATCTGTCAGATGTCAGAGAGAGAGAGAGAGAGAGAGAG	1720	
OY	1506	TATACATGCTGACCAATTAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1565	
Db	1721	TATACATGCTGACCAATTAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1780	
OY	1566	ACCTGGAG	1625	
Db	1781	TGTGTGCTGCTGACCAATTAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1840	
OY	1626	CTGAGCATGAG	1685	
Db	1841	CTGAGCATGAG	1897	
OY	1686	TCAGGCTGTCAGATCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1745	
Db	1898	CCTGATGCTTATTCATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1957	
OY	1746	TCTTCTGCTGCTGCTGATGACATATCATGACCTTGTGAAAGAGAGAGAGAGAG	1805	
Db	1958	TCTTCTGCTGCTGCTGATGACATATCATGACCTTGTGAAAGAGAGAGAGAGAG	2017	
OY	1806	TGGAATGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1865	
Db	2018	TGGAATGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2077	
OY	1866	ACAAAGCCCGGAG	1925	
Db	2078	GAGCAG	2137	
OY	1926	ACTTTCACCTTGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1982	
Db	2138	ACATTCACATGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	2197	
OY	1983	CCATACACCAAGCAG	2042	
Db	2198	CCCTACAG	2257	
OY	2043	ATCATGATGAG	2102	
Db	2258	GTCATGAGTGGAG	2317	
OY	2103	AAG	2129	
Db	2318	AAAG	2344	

RESULT 12
B1108275 719 bp mRNA EST 26-JUN-2001
LOCUS

DEFINITION 602902137f1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5032016 5', mRNA sequence.
ACCESSION B1108275
VERSION B1108275.1 GI:14559181
KEYWORDS EST
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 719)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM1089 row: 1 column: 09
High quality sequence stop: 719.
Location/Qualifiers
1. 719
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:5032016"
/clone_lib="NIH_CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert 2 kb. Library constructed by Life Technologies, catalog #12017-018. Investigators providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999). Note: this is a NCI_CGAP Library."
BASE COUNT 184 a 198 c 187 g 150 t
ORIGIN
Query Match 24.2%; Score 694.4; DB 11; Length 719;
Best Local Similarity 99.9%; Pred. No. 6,2e-149;
Matches 693; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1679 GAAGTACTGAGGTGTGACATGATGCTAAATCTGCAAGAAACATGCTGCGAA 1738
DB 24 GAAGTACTGAGGTGTGACATGATGCTAAATCTGCAAGAAACATGCTGCGAA 83
OY 1739 GGGCTTCTCTTCTGGGCTGCTGCTAGACAAATATCATGACCTTGTGAAAAAGTATATCTT 1798
DB 84 GGGCTTCTCTTCTGGGCTGCTGCTAGACAAATATCATGACCTTGTGAAAAAGTATATCTT 143
OY 1799 GGGCTTCTCTTCTGGGCTGCTGCTAGACAAATATCATGACCTTGTGAAAAAGTATATCTT 1798
DB 144 GGGCTTCTCTTCTGGGCTGCTGCTAGACAAATATCATGACCTTGTGAAAAAGTATATCTT 203
OY 1859 CCTAAGCAACAAAGCCCGGGGACCTTCTACTGCGCTTACAGAGAGAGAGCAAGCAAG 1918
DB 204 CCTAAGCAACAAAGCCCGGGGACCTTCTACTGCGCTTACAGAGAGAGAGCAAGCAAG 263
OY 1919 AGGGGTCACTTTCATCTGGGTGGAAAAAGACATCACTGAGCAAGAGCCAGATCCAGTCTGT 1978
DB 264 AGGGGTCACTTTCATCTGGGTGGAAAAAGACATCACTGAGCAAGAGCCAGATCCAGTCTGT 323
OY 1979 AGAGGCATACACCAAG 2038
DB 324 AGAGGCATACACCAAG 383
OY 2039 TAAGTCAATGATGAGGACCAACATCTGTGCTCAGTGTCTACCTACCCGACAT 2098
|||||

DB 384 TAAGTCAATGATGAGGACCAACATCTGTGCTCAGTGTCTACCTACCCGACAT 443
OY 2099 TCCCAAG 2158
DB 444 TCCCAAG 503
OY 2159 CGAGCCAGGAGTGTGTCGCCCTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAG 2218
DB 504 CGAGCCAGGAGTGTGTCGCCCTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAG 563
OY 2219 CTCGACCAATACATGATGAGTGTGTCGCCCTGAGTGTGAGAGAGAGAGAGAG 2278
DB 564 CTCGACCAATACATGATGAGTGTGTCGCCCTGAGTGTGAGAGAGAGAGAGAG 623
OY 2279 TCGAATTAACGAGTGTGTCGCCCTGAGTGTGAGAGAGAGAGAGAGAGAGAGAG 2338
DB 624 TCGAATTAACGAGTGTGTCGCCCTGAGTGTGAGAGAGAGAGAGAGAGAGAGAG 683
OY 2339 TCGAATTAACGAGTGTGTCGCCCTGAGTGTGAGAGAGAGAGAGAGAGAGAGAG 2374
DB 684 TCGAATTAACGAGTGTGTCGCCCTGAGTGTGAGAGAGAGAGAGAGAGAGAGAG 719
RESULT 13
BG973911 714 bp mRNA EST 12-JUN-2001
LOCUS 602843546f1 NIH_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4979458 5', mRNA sequence.
DEFINITION BG973911
ACCESSION BG973911
VERSION BG973911.1 GI:14361548
KEYWORDS EST
SOURCE house mouse.
ORGANISM Mus m. musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 714)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM1097 row: k column: 11
High quality sequence stop: 714.
Location/Qualifiers
1. 714
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:4979458"
/clone_lib="NIH_CGAP_Mam4"
/tissue_type="tumor, gross tissue"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert 2.5 kb. Library constructed by Life Technologies, catalog # 12018-016. Investigators providing samples: Lothar Hennighausen/Priscilla Furth, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996). Note: this is a NCI_CGAP Library."
BASE COUNT 184 a 187 c 195 g 148 t
ORIGIN
Query Match 24.0%; Score 688.4; DB 11; Length 714;
Best Local Similarity 99.6%; Pred. No. 1.5e-147;
|||||

Matches 711: Conservative 0; Mismatches 1; Indels 2; Gaps 2;

```

Oy 1284 GCAGAGTTCAACGACCTGACCCCTTAGGAGCAGAGATGTGGATGAGAGCCGCTGCCAAT 1343
      |||||||
Db 1 GCAGAGTTCAACGACCTGACCCCTTAGGAGCAGAGATGTGGATGAGAGCCGCTGCCAAT 60

Oy 1344 TGTGATGCTCTCTTGTGCTGACTGAGAGCTGCACCTGATCATCTTCGAGACTGAGGTG 1403
      |||||||
Db 61 TGTGATGCC TCTGTATGCTGACTGAGAGCTGCACCTGATCATCTTCGAGACTGAGGTG 120

Oy 1404 TGCACCAAGGCGCTCAAGATTGACCTAGAGACCCATCTCTGTCAGTTGGTGTGATCTGC 1463
      |||||||
Db 121 TGCACCAAGGCGCTCAAGATTGACCTAGAGACCCATCTCTGTCAGTTGGTGTGATCTGC 180

Oy 1464 AACATCTGTACATGCCAATATGCTTGGGATCAATCCTGTATTAACATGCTGACCAAT 1523
      |||||||
Db 181 AACATCTGTACATGCCAATATGCTTGGGATCAATCCTGTATTAACATGCTGACCAAT 240

Oy 1524 AACCCCAAGACGTGAACCTCTGACTAAGCCGCAATTGGAACTGGGAGCAAGTGGCC 1583
      |||||||
Db 241 AACCCCAAGACGTGAACCTCTGACTAAGCCGCAATTGGAACTGGGAGCAAGTGGCC 300

Oy 1584 GAGGTGCTGAGCTGGCAGTTCTCG -TCCACCAAGCAGGCGCTGAGCATCGAG-AGCT 1642
      |||||||
Db 301 GAGGTGCTGAGCTGGCAGTTCTCGTCCACCAAGCGGGGCTGAGCATCGAGAGCT 360

Oy 1643 GACACGCTGCTGAGACAGCTCCTAGGCGCTGCTGTGAACTACTCAAGGTGTAGATCAC 1702
      |||||||
Db 361 GACACGCTGCTGAGACAGCTCCTAGGCGCTGCTGTGAACTACTCAAGGTGTAGATCAC 420

Oy 1703 ATGGGCTAAATTTGCAAGAAACATGCTGCGAAGGCGTTCTCTTCTGGGCTTGCGT 1762
      |||||||
Db 421 ATGGGCTAAATTTGCAAGAAACATGCTGCGAAGGCGTTCTCTTCTGGGCTTGCGT 480

Oy 1763 AGCAATATTCATGACCTTGTGAAGAAATATCTTGGCCCTTTGGAAATGAAGGATCAT 1822
      |||||||
Db 481 AGCAATATTCATGACCTTGTGAAGAAATATCTTGGCCCTTTGGAAATGAAGGATCAT 540

Oy 1823 CATGGGTTTCATGACAGAGAGGCGGCGGCGATCTCAAGCAAAAGCCCGGGCAC 1882
      |||||||
Db 541 CATGGGTTTCATGACAGAGAGGCGGCGGCGATCTCAAGCAAAAGCCCGGGCAC 600

Oy 1883 CTTCCTACTGCGCTTGACGAGAGCAAGCAAGAGGCGTCACTTTCAGTTGGGTGA 1942
      |||||||
Db 601 CTTCCTACTGCGCTTGACGAGAGCAAGCAAGAGGCGTCACTTTCAGTTGGGTGA 660

Oy 1943 A-AAGGACATCAGTGGCAGACCCAGATCCAGTCTGTAGAGCCATACACCAAGC 1995
      |||||||
Db 661 ACAAGGACATCAGTGGCAGACCCAGATCCAGTCTGTAGAGCCATACACCAAGC 714

```

RESULT 14

B1155944 736 bp mRNA EST 05-JUL-2001

LOCUS 602904278f1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:503395 5',

DERIVATION mRNA sequence.

ACCESSION B1155944

VERSION B1155944.1 GI:14615957

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE NIH-MGC http://mgi.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Lotmar Hennighausen Ph.D., Chu-Xia Deng Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLAM1094 row: k column: 20

High quality sequence stop: 732.

FEATURES

Location/Qualifiers

```

1..736
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="NIH_CGAP_Mam3"
/issue_type="tumor, gross tissue"
/ab_host="DH10B"
/notes="Organ: mammary; Vector: pCMV-Sport6; Site:1; Not1; Site:2; Sal1; Cloned unidirectionally. Primer: Oligo dT. Average insert 2 kb. Library constructed by Life Technologies, catalog #12017-018. Investigators providing samples: Lotmar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999). Note: this is a NCI_CGAP Library."
BASE COUNT 199 a 178 c 204 g 155 t
ORIGIN

```

Query Match 23.7%; Score 679; DB 11; Length 736;

Best Local Similarity 98.8%; Pred. No. 2.1e-145;

Matches 726: Conservative 0; Mismatches 5; Indels 4; Gaps 4;

```

Oy 735 GGGCTCTTGCACCAATGAGTACGTCGACAGACACTGACTGATGAAGAGCTGGCTGAC 794
      |||||||
Db 1 GGGCTCTTGCACCAATGAGTACGTCGACAGACACTGACTGATGAAGAGCTGGCTGAC 60

Oy 795 TGGAGAGGGGGGCGAGATCGGTCGATCGAGGCGCTCCCAACATCTGCTGGACCGT 854
      |||||||
Db 61 TGGAGAGCGCGGCGAGATCGGTCGATCGAGGCGCTCCCAACATCTGCTGGACCGT 120

Oy 855 CTGAA--TCTGTATACTTCATTAGCAGATCTCAACTTCAGACCCGCCAACAATTAA 913
      |||||||
Db 121 CTGAACTGCTGATACCTTCATTAGCAGATCTCAACTTCAGACCCGCCAACAATTAA 180

Oy 914 GAAACTGGAAGAGCTGAGGAAAGTGTCTCAAGGGGAGCCCTATCTGCGACGACCG 973
      |||||||
Db 181 GAAACTGGAAGAGCTGAGGAAAGTGTCTCAAGGGGAGCCCTATCTGCGACGACCG 240

Oy 974 GCCCATGCTGAGAGAGAGATCTGTGAGCTGTTCAGAACTTAATGAAGAGTCCCTTGT 1033
      |||||||
Db 241 GCCCATGCTGAGAGAGAGATCTGTGAGCTGTTCAGAACTTAATGAAGAGTCCCTTGT 300

Oy 1034 GTTGAGGCGGAGCCCTTCGATGCCATGACCCGAGCCGCCCTTACTCATCAAGACTGG 1093
      |||||||
Db 301 GTTGAGGCGGAGCCCTTCGATGCCATGACCCGAGCCGCCCTTACTCATCAAGACTGG 360

Oy 1094 TGTTCATTACACAGAAACTGAGTGTGCTGTCGAATTTCTGAGTTGATATGAGCT 1153
      |||||||
Db 361 TGTTCATTACACAGAAACTGAGTGTGCTGTCGAATTTCTGAGTTGATATGAGCT 420

Oy 1154 TAAATTTAAAGTGTGATTTGAAGAGCTGGGAGTGTTCGCTCAGAGGCTTCG 1213
      |||||||
Db 421 TAAATTTAAAGTGTGATTTGAAGAGCTGGGAGTGTTCGCTCAGAGGCTTCG 480

Oy 1214 G-AAATTTAATCTTGGGCGACGACAAAGAGTATGAACATGAGAGATCTAACAGC 1272
      |||||||
Db 481 GAAATTTAATCTTGGGCGACGACAAAGAGTATGAACATGAGAGATCTAACAGC 540

Oy 1273 GCAAGCTGTCTGC-AGAGTTCAAGCACCTGACCTTGAAGAGCAGAGATGTGGAATGA 1331
      |||||||
Db 541 GCAAGCTGTCTGC-AGAGTTCAAGCACCTGACCTTGAAGAGCAGAGATGTGGAATGA 600

Oy 1332 GGGCGTCCCAATTTGATGGCTCTTGTATGCTGACTAGAGAGCTGCACCTGATCCTTC 1391
      |||||||
Db 601 GGGCGTCCCAATTTGATGGCTCTTGTATGCTGACTAGAGAGCTGCACCTGATCCTTC 660

Oy 1392 GAGACTGAGTGTACCAAGGCGCTCAAGATTGAGCT-AGAGACCCACTCTTGGCCAGT 1450

```

```
Db 661 GAGACTGGGCTGACCAAGGCGCTCAAGATGACCTAAGAGACCACTCTCCAGT 720
OY 1451 TGTGCTGATCTCCAA 1465
Db 721 AGTGCTGATCTCCAA 735

RESULT 15
B1113023 717 bp mRNA EST 26-JUN-2001
LOCUS 602898184F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5028140 5',
DEFINITION mRNA sequence.
ACCESSION B1113023 GI:14563924
VERSION B1113023.1
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 717)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9qpb5-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11079 row: g column: 21
High quality sequence stop: 678.
Location/Qualifiers
1..717
/organism="Mus musculus"
/strain="C57/B6"
/db_xref="taxon:10090"
/clone="IMAGE:5028140"
/clone_lib="NCI_CGAP_Mam5"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"
BASE COUNT 188 a 193 c 192 g 144 t
ORIGIN

Query Match 23.6%; Score 676.6; DB 11; Length 717;
Best Local Similarity 99.2%; Pred. No. 7.6e-145;
Matches 701; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

OY 1536 GTGACCTCTTCAAGCGCCCAATTGGAACCTGGACCAAGTGGCCGAGGCTCAGC 1595
Db 1 GTGAACCTCTTCAAGCGCCCAATTGGAACCTGGACCAAGTGGCCGAGGCTCAGC 60
OY 1596 TGGCAGTTCTCGTCGACCAAGGAGGCTGAGCATGAGCAGCTGACGTCAGC 1655
Db 61 TGGCAGTTCTCGTCGACCAAGGAGGCTGAGCATGAGCAGCTGACGTCAGC 120
OY 1656 GAGAGCTCCTAGGCGCTGGTGAATCACTAGGCTGTCAGATCAGATGGCTAAATTC 1715
Db 121 GAGAGCTCCTAGGCGCTGGTGAATCACTAGGCTGTCAGATCAGATGGCTAAATTC 180
OY 1716 TGCAAGAGAAACATGCTGGCAAGGCTTCTCTCTGAGCTGGCTAGACAAATATCATC 1775
Db 181 TGCAAGAGAAACATGCTGGCAAGGCTTCTCTCTGAGCTGGCTAGACAAATATCATC 240
```

```
OY 1776 GACCTTGTGAAAAGTATATCTTGCCCTTTGGATGAAGGCTACATCATGGTTTCATC 1835
Db 241 GACCTTGTGAAAAGTATATCTTGCCCTTTGGATGAAGGCTACATCATGGTTTCATC 300
OY 1836 AGCAAGGAGCGGAGCGGCGCATCTTAAGCACAAAGCCCCGGGACCTCTCTACTGCGC 1895
Db 301 AGCAAGGAGCGGAGCGGCGCATCTTAAGCACAAAGCCCCGGGACCTCTCTACTGCGC 360
OY 1896 TTCACGAGAGCAGCAAGCAAGAGGAGGTCACCTTTCACCTTGGGTGAAAAGACATCAGT 1955
Db 361 TTCACGAGAGCAGCAAGCAAGAGGAGGTCACCTTTCACCTTGGGTGAAAAGACATCAGT 420
OY 1956 GGCAGACCCAGATCCAGCTGTGTAGACCATACAGCAGCAGCAGCTGTAACAACATGCA 2015
Db 421 GGCAGACCCAGATCCAGCTGTGTAGACCATACAGCAGCAGCAGCTGTAACAACATGCA 480
OY 2016 TTTGCTGAATCATCATGGCTATTAAGATCAT -GGATGCGACCAACATCTGTGTCTCC 2074
Db 481 TTTGCTGAATCATCATGGCTATTAAGATCATGGATGGACCAACATCTGTGTCTCC 540
OY 2075 ACTTGCTACCTCTACCCCGACATTTCCCAAGGAGAGGATTTGGAAGTACTGAGGCC 2134
Db 541 ACTTGCTACCTCTACCCCGACATTTCCCAAGGAGAGGATTTGGAAGTACTGAGGCC 600
OY 2135 CGAGAGCTCAGAGCACCCCGAAGCCGACCCAGTAGTGCTGCCCGTACCTGAAGACCAA 2194
Db 601 CGAGAGCTCAGAGCACCCCGAAGCCGACCCAGTAGTGCTGCCCGTACCTGAAGACCAA 660
OY 2195 GTTCATCTGTGTGACACCAAGCAGCTGCAAGCAATACATTGACCTGC 2241
Db 661 GTTCATCTGTGTGACACCAAGCA -CTGCGCAATTAACATTGACCTGC 706
```

Search completed: March 19, 2002, 14:30:29
Job time: 1772.0000


```
QY 301 HRPMLEERIVELFRNLKKSAPVVEROPCMHPDRPLVKTGVQFTTKVLLYKPELNT 360
|||||
Db 301 HRPMLEERIVELFRNLKKSAPVVEROPCMHPDRPLVKTGVQFTTKVLLYKPELNT 360
QY 361 QLKTKVCIDKDSGVAALRGSRKFNILGTNTKVMNEESNNGSLSAEFKHLTLREORCGN 420
|||||
Db 361 QLKTKVCIDKDSGVAALRGSRKFNILGTNTKVMNEESNNGSLSAEFKHLTLREORCGN 420
QY 421 GGRANCASLIIVTEELHLITFEFEVYHQGLKIDLETHSLPVVVISNICOMPMAASILMY 480
|||||
Db 421 GGRANCASLIIVTEELHLITFEFEVYHQGLKIDLETHSLPVVVISNICOMPMAASILMY 480
QY 481 NMLTNKRNKNVFTKPP1GTMDOVAEVLWSQFSTTKRGSLIEOLTLAKELLCGPNYS 540
|||||
Db 481 NMLTNKRNKNVFTKPP1GTMDOVAEVLWSQFSTTKRGSLIEOLTLAKELLCGPNYS 540
QY 541 GCOTTMKFCCKENNAKGFSFWMLDNIIDLKYYILALMNEGYIMGFISKEBERALIST 600
|||||
Db 541 GCOTTMKFCCKENNAKGFSFWMLDNIIDLKYYILALMNEGYIMGFISKEBERALIST 600
QY 601 KPCTFLFRSESSKEGCVFTFWWEKDISGKTQIOSVEPYTKOOLNNMFAEIIIMGYKIM 660
|||||
Db 601 KPCTFLFRSESSKEGCVFTFWWEKDISGKTQIOSVEPYTKOOLNNMFAEIIIMGYKIM 660
QY 661 DATNILVSPLYLYLPDIPKEEAFGKYCRPESQEHPEADPGSAARYLTKFICVPTTCGN 720
|||||
Db 661 DATNILVSPLYLYLPDIPKEEAFGKYCRPESQEHPEADPGSAARYLTKFICVPTTCGN 720
QY 721 TIDLPMSPRTLDSLMQFGNNGEAPSGAGOFESLTFDMDLTSECATSPM 770
|||||
Db 721 TIDLPMSPRTLDSLMQFGNNGEAPSGAGOFESLTFDMDLTSECATSPM 770

RESULT 2
DNA-binding protein APRF - human
C:Species: Homo sapiens (man)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 28-Jul-2000
C:Accession: A54444
R:Akira, S.; Nishio, Y.; Inoue, M.; Wang, X.J.; Wei, S.; Matsusaka, T.; Yoshida, K.; Sue
Cell 77, 63-71, 1994
A:Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related tra
A:Reference number: A54444; MUID:94208062
A:Accession: A54444
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-770 <RES>
A:Cross-references: GB:I29277; NID:g475788; PID:g475789
C:Genetics:
A:Gene: GDB:STAT3; APRF
A:Cross-references: GDB:358950
A:Map position: 17q21-17q21
C:Superfamily: human signal transducer and transcription activator STAT5A
C:Keywords: DNA binding; transcription factor
```

```
Query Match 98.7%; Score 3977; DB 2: Length 770;
Best Local Similarity 98.6%; Pred. No. 7,le-238;
Matches 759; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 MAONMLOOUDTRYLKOLHOLYSDFPMELOFLAPWIESODMAVAASKESHATLVFHNL 60
|||||
Db 1 MAONMLOOUDTRYLKOLHOLYSDFPMELOFLAPWIESODMAVAASKESHATLVFHNL 60
QY 61 LGEIDQOYSRFLQESNVLYOHNLRRIKQFLOSRYLEKPMETIARIIVARCIMESRLIQTAA 120
|||||
Db 61 LGEIDQOYSRFLQESNVLYOHNLRRIKQFLOSRYLEKPMETIARIIVARCIMESRLIQTAA 120
QY 121 TAAOOGGANHTAAYVTEKOQMLEOHLODVRRKRVODLEOKMKVVENLODDDFENYKTLK 180
|||||
Db 121 TAAOOGGANHTAAYVTEKOQMLEOHLODVRRKRVODLEOKMKVVENLODDDFENYKTLK 180
```

```
QY 181 SOGDMODLNNNSVTRQKMOOLEOMITALDOMRSTIVSELAGLLSMEVYOKTLTDEEL 240
|||||
Db 181 SOGDMODLNNNSVTRQKMOOLEOMITALDOMRSTIVSELAGLLSMEVYOKTLTDEEL 240
QY 241 ADMKRRPEIACIGPPN1CIDLLENMTTSLAESOLQTRQOIKKLEELOOKVSKGDP1VQ 300
|||||
Db 241 ADMKRRQOIAICIGPPN1CIDLLENMTTSLAESOLQTRQOIKKLEELOOKVSKGDP1VQ 300
QY 301 HRPMLEERIVELFRNLKKSAPVVEROPCMHPDRPLVKTGVQFTTKVLLYKPELNT 360
|||||
Db 301 HRPMLEERIVELFRNLKKSAPVVEROPCMHPDRPLVKTGVQFTTKVLLYKPELNT 360
QY 361 QLKTKVCIDKDSGVAALRGSRKFNILGTNTKVMNEESNNGSLSAEFKHLTLREORCGN 420
|||||
Db 361 QLKTKVCIDKDSGVAALRGSRKFNILGTNTKVMNEESNNGSLSAEFKHLTLREORCGN 420
QY 421 GGRANCASLIIVTEELHLITFEFEVYHQGLKIDLETHSLPVVVISNICOMPMAASILMY 480
|||||
Db 421 GGRANCASLIIVTEELHLITFEFEVYHQGLKIDLETHSLPVVVISNICOMPMAASILMY 480
QY 481 NMLTNKRNKNVFTKPP1GTMDOVAEVLWSQFSTTKRGSLIEOLTLAKELLCGPNYS 540
|||||
Db 481 NMLTNKRNKNVFTKPP1GTMDOVAEVLWSQFSTTKRGSLIEOLTLAKELLCGPNYS 540
QY 541 GCOTTMKFCCKENNAKGFSFWMLDNIIDLKYYILALMNEGYIMGFISKEBERALIST 600
|||||
Db 541 GCOTTMKFCCKENNAKGFSFWMLDNIIDLKYYILALMNEGYIMGFISKEBERALIST 600
QY 601 KPCTFLFRSESSKEGCVFTFWWEKDISGKTQIOSVEPYTKOOLNNMFAEIIIMGYKIM 660
|||||
Db 601 KPCTFLFRSESSKEGCVFTFWWEKDISGKTQIOSVEPYTKOOLNNMFAEIIIMGYKIM 660
QY 661 DATNILVSPLYLYLPDIPKEEAFGKYCRPESQEHPEADPGSAARYLTKFICVPTTCGN 720
|||||
Db 661 DATNILVSPLYLYLPDIPKEEAFGKYCRPESQEHPEADPGSAARYLTKFICVPTTCGN 720
QY 721 TIDLPMSPRTLDSLMQFGNNGEAPSGAGOFESLTFDMDLTSECATSPM 770
|||||
Db 721 TIDLPMSPRTLDSLMQFGNNGEAPSGAGOFESLTFDMDLTSECATSPM 770
```

```
RESULT 3
A46159
interferon-dependent positive-acting transcription factor ISGF-3 91k chain - human
C:Species: Homo sapiens (man)
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 28-Jul-2000
C:Accession: A46159
R:Schindler, C.; Fu, X.Y.; Impirota, T.; Aebersold, R.; Darnell Jr., J.E.
Proc. Natl. Acad. Sci. U.S.A. 89, 7836-7839, 1992
A:Title: Proteins of transcription factor ISGF-3: one gene encodes the 91- and 84-kDa
A:Reference number: A46159; MUID:92366557
A:Accession: A46159
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid; protein
A:Residues: 1-739 <SCH>
A:Experimental source: HeLa cells
A>Note: Sequence extracted from NCBI backbone (NCBIP:110818)
C:Superfamily: human signal transducer and transcription activator STAT5A
```

```
Query Match 48.9%; Score 1971.5; DB 2: Length 739;
Best Local Similarity 51.7%; Pred. No. 4,7e-114;
Matches 383; Conservative 144; Mismatches 189; Indels 25; Gaps 10;
```

```
QY 1 MAONMLOOUDTRYLKOLHOLYSDFPMELOFLAPWIESODMAVAASKESHATLVFHNL 60
|||||
Db 1 MSOWEYLOOUDSKLEOVHOLYDSSFPMETROYLAOWLEKODMEHAAANDVSFAIRFHD 60
QY 61 LGEIDQOYSRFLQESNVLYOHNLRRIKQFLOSRYLEKPMETIARIIVARCIMESRLIQTAA 120
|||||
Db 61 LSQLDQOYSRFLQESNVLYOHNLRRIKQFLOSRYLEKPMETIARIIVARCIMESRLIQTAA 120
QY 121 TAAOOGGANHTAAYVTEKOQMLEOHLODVRRKRVODLEOKMKVVENLODDDFENYKTLK 180
|||||
Db 121 TAAOOGGANHTAAYVTEKOQMLEOHLODVRRKRVODLEOKMKVVENLODDDFENYKTLK 180
```


A:Reference number: S53873; MUID:95192056
A:Accession: S53873
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-196:392-591:684-730 <YAM>
A:Cross-references: EMBL:U18671
C:Genetics:
A:Gene: stat2
A:Introns: 44/2; 95/3; 127/3; 157/3; 183/1; 211/3; 261/2; 314/2; 345/2; 365/2; 372/2; 40
C:Superfamily: human signal transducer and transcription activator STAT5A
C:Keywords: signal transduction; transcription regulation

Query Match 32.1%; Score 1293; DB 2; Length 851;
Best Local Similarity 38.6%; Pred. No. 4.6e-72;
Matches 298; Conservative 148; Mismatches 276; Indels 50; Gaps 18;
Db 1 MAOMNLOOLDTRYLKLHQLYSDF-FPMELROFLAPMIESODMAVAA--SKESHATLV 57
1 MAOMENLOLDSPFODQLHQLYSHSLRPVDIRQYLAVMTEEDQNMGEALGSDSKATHLF 60
58 HMLGEIDQOYSRFLQF--SNVLYOHNLRIKOFLOSRYLEKPMETARIYARCLMEESRL 116
61 FHFLODLNECCRCGSDPELILQNLKRFCDIP-FSODPTOLAEMLFNLLLEKRL 119
117 QTAATAAGGGOANHPRTAAVTEKOO-MLEQHLODVRKRVODLEQMKVEMLODDPFN 175
120 IOAORAOLEQF---PVELETPEVSOOHEESRIIDLRAEMELVKSISQLKQODVFCFR 176
176 YVTLKSQGMODLNGNNSVTRKMOOLEOMLTALDOMRSTIVSELACLILSMEYVOKTL 235
177 YK-IOAKGKTPSLDPH--OTKEOKI--LOETLNEIDRKREVLDAISKALGLTTLIEL 231
236 TDEELADMKRREBICIGGPPNICDLRENMWITSLSAESQLOTROOKKLELOOKVSYK 295
232 L-PKEEMKAQOQKACIRAPIDHGLEOLEFTWTAGAKLLFHLROLLEKELGSLCVSYOD 290
296 DPVQHRPMLERIVLEFRLNLSAFVVEROPCMHPDRPLVYIKTGVQFTTKVRLLYKF 355
291 DPLTKQVDLRNAQVTELLORLHLRAFVVEVETOPCMQPTPRPLILKGSFTVTRLLVRL 350
356 PELNVOGLKIKVCIDKSDGVAALRGSRKFNIIIGTNTKVMNMESSNNGSLSAEKLHLTRE 415
351 OEGNESLTVESIDRNP---POLQFRKRNILTSNOKTLTPREKOSQGLIMDFGYTLVE 407
416 ORCGNGRANCASLIVTEELHLITFETEVYHQGLKIDLETSLPVVVISNCOMPNAA 475
408 QREGSGGKSGNKPGLVTEELHIISTVAKYTYOGLKQELKTDTLPVVIISNMQSLIAMA 467
476 SILMYMMLNRRKNVFFTRKPIGTMDOVAEVLWSQFSTTKRGSLIDQUTLAEKLLCP 535
468 SYLWFLNL-LPNLONOGFESNPAPASLGLPALMSQFSSYVGGRLSDSLSRNKLRFQ 527
536 GVVYSQCOITMAFECK-ENMAGKGFSEFWMLDNIIDLVRKYILLANEGLMGFISKEBE 594
528 NCTEDPRLSMADFTKRESPPCK-LPFWIMLOKILIELVHDHLKLDNDRINGFVSRSBE 586
595 RAILSTRKPGCTFLRFSSESGEGVTFWVEKDISKTOIOSVEPTTKOOLNMNMAETII 654
587 RRLKTKMSTGFLRFSSESS-EGGICISWVEHODDKVLIVSYQPTKTEVLOSPLTEII 645
655 MGKIKIDATITLIVSPLYVLYPOIPKEEAFGKYCRPE--SOEHPEAPGSAAPLAKFKFC 712
646 RHQQLTEENIPENPLRFLYPRIPROEAFGQYQEKVNIQERKK-----YLKRLILV 697
713 VPTTCSNTIDLPMSPTLDSLMQFGNNGEAPSGAGFESITFDMDLTSE 764
698 V-----SNROYDELQO-----PLELKPEPELESLELGLVPE 730

RESULT 6
G02317
transcription activator stat5a - human

C:Species: Homo sapiens (man)
C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 28-Jul-2000
C:Accession: G02317
R:Lin, J.
submitted to the EMBL Data Library, December 1995
A:Reference number: H01043
A:Accession: G02317
A:Status: preliminary; translated from cDNA/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-794 <LIN>
A:Cross-references: EMBL:U043185; NID:q1151169; PIDN:AAB06589.1; PID:q1151170
C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 21.6%; Score 869.5; DB 2; Length 794;
Best Local Similarity 30.5%; Pred. No. 5.9e-46;
Matches 249; Conservative 146; Mismatches 315; Indels 107; Gaps 29;
Db 1 MAOMNLOOLDTRYLKLHQLYSDF-FPMELROFLAPMIESODMAVAA--ASKESHATLV 56
1 MAOMTQAOQLOGDALRQMQVLYGQHFPLEVRHYLAOMTIESQPMDAIDILNPQDAQTOL 60
57 FHNILGEIDQOYSRFLQF--SNVLYOHNLRIKOFLOSRYLEKPMETARIYARCLMEESRL 116
61 LEGIAVQELQKRAEHQVGDGFLIKIKLKHVATVQDTYDKRCLVLRCKIKHILYDEQVILV 120
117 QTAATAAGGGOANHPRTAAVTEKOO-MLEQHLODVRKRVODLEQMKVEMLODDPFN 176
121 REANNCSSPAG-----ILVDAMSKHLOJINUFFEELRLVTOJTEENELKKLOOTGEYFIIY 176
177 KTLKSQGMODL-----NGNNSVTRKMOOLEOML-----TALDOMRSTIVSELACL 224
177 QESLRIOAFOALQSLDSPOERLSRETALQOKVSLSEMLEOREAOTLOQRYVLEAEKHQKT 236
225 LSAMEYVOKTLTDEELADMKRREBICIGGPPNICDLRENMWITSLSAESQLOTROOKKLE 284
237 LQLLKQOOTIILDDDELIOKKRQOQLAGNGRPFESSLDVLOSCKELAEIIMQNRQILRA 296
285 BELQOKVSYKQDPIVQHRPMLERIVLEFRLNLSAFVVEROPCMHPDRPLVYIKTGVQ 344
297 EHICQOLPIPG-PVEEMLAENVATITDIISALVSTFIIIEKOP-----POVLKTYTK 347
345 FTTKRYRLLVKPEELNVOL---KIKVCIDKSDGVAALRGSRKFN---IIGTNTKVMN 396
348 FAATVRLVVG-GKLNVHNMPPQVKATITISEOASSLNENMTREDCSEL-NNCCMEY 405
397 EESNNGSLSAEKLHLTREORCGNGRANCASLIVTEELHLITFETE--VYHGLKIDL 454
406 HQA-TGTLISAHFRNMSLRKIK-----RADRGAESVTEEEKFTVLFEQFVSQNELVFOV 459
455 ETHSLPVPVVISNCOMPNANASILMYNMLJNRPKNVNFTRKPIGTMDOVAEVLWSQFSS 514
460 KTLSPVPVVIYVHSGSDHNAATVLDNDAFA-EPGRVP-FAVPDVKVLPQLOCEALNMFKFA 517
515 --TTRKGLSIEQUTLAEKLLPG---VNYSGCOITMAFKCKENMAGKGFSEFWMLDNI 568
518 EYOSNRKGLTKENLVLAQKLFNNSSSHLEDYSGLSVSSQGNKRELPMWNTTFQWMDGV 577
569 IDLVVKYILLANEGLMGFISKEKERAILSTRKPGCTFLRFSSESGEGVTFWVEKDI 628
578 MEVLKHHKHPHWNDAIILGFVNKQOAHDLINKPDPCTFLRFS-SEIGGITTAM--KFD 634
629 SGKTOIOSVEPTTKOOLNMNMAETIIMGYKIMIDATNIIIVSPLYVLYPDIPKEEAFGKYCR 688
635 SPERILMNLKPTTIDFTRSLAD-----RLQD-----LSLTIIVFPDRPDEVFSKTYT 684
689 PE-----SOEHPE-----ADP-GSAAPYLKTKFCVYPTTCSNTIDLPMSPR 729
685 PVLAKAVDQYVQKQYVVEFVNASADAGSSATYMDQ---ASBPACPOA-PYNNYMQ 740
730 TIDSLMQFGNNGEAPSGAGFESITFDMDLTSCFA 766
741 NPDHYILD--QDGE-----FDLDETMDVA 761

Db 406 HQA--TGTLTAIFRHMMLSLKRIKRSDBRGAGS-----VTEKKFTLLPDSQFSVGNEL 455

Oy 451 KIDLETSLPVPVYINSLICOMPMANASIIIMYMLTLNPKNNFFTKPKPIGWIQYAEVLSW 510

Db 456 VFQVKTSLPVPVYIVHGSQDNNATATATVIMDAFA-EGRAVP-FVAPKVLMPQLCEALNM 513

Oy 511 QESS--TTRKRLSIEOLUTLAETKLIGPGVN-----YSCQQLTWAFCKENMAGKFSFVWV 564

Db 514 KFKAEVQSNRCLTENI.VFLAQKLFNISSNILEYNSKMSVMSQFNPHNLPGRAYTWQW 574

Oy 565 LDNIIDLVKKYILALMNEGYMGFISKEEREAIIISTKPPGFTLLRPSSESKGGVFTTWV 624

Db 574 FDGVMEVYIKHLKPLIMNDGAILIGFVNKKQADHLLINKPDPFELLRFSD-SEIGGITIAM- 631

Oy 625 EKDISGKTIQSVVEPYTQOILNNMSPAFIIMGYIMATNIIIVSPLVLYPDIIPKEAFG 684

Db 632 -KFDQSEHMYMLPFTTKDPSIRSLAD-----RLGD-----LWYLLIYFPDKRKDEYIS 680

Oy 685 KY-----CRPES-----QEHP-----ADPGSAAPYLKTKFICVPTPTGSNT 721

Db 681 KYTTVPCEPATAAAGGYVKKQIKQVVPPEPANAStAGATVMDQ---APSPVCFQIA 737

Oy 722 IDLPMSPTLLSLMFGNNNGCAERPSAGGQFESLTFFMDLTISECA 766

Db 738 -HYNMPYPPDSVLDD--TDGD-----FDLEDYIMIVA 765

RESULT 8

SS4772

mammary gland factor - mouse

N:Alternate names: stat5 protein

C:Species: Mus musculus (house mouse)

C:Date: 27-Oct-1995, sequence_revision 03-Nov-1995 #1:vt_change 28-Jun-2000

C:Accession: S54772; 149273

R:Mu, A.L.F., M., H., J'Farrell, A.M., Harada, N., Miyajima, A.

EMBO J. 14, 1166-1175, 1995

A>Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin-3

A:Reference number: S54772; MUID:95237198

A:Accession: S54772

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-772

A:Cross-reference: EMBL:Z48538; NID:9758633; PIDD:CA88419.1; PID:9758634

R:Li, X.; Robinson, G.W.; Gouilleux, F.; Groner, R.; Hennighausen, L.

Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995

A>Title: Cloning and expression of Stat5 and an α -

A:Reference number: 119273; MUID:96004632

A:Accession:

A:Status: p

A:Molecule type:

A:Residues:

A:Cross-reference: MBL:U21103; NID:9747971; PII:

C:Genetics:

A:Gene: Stat5

C:Superfamily:

signal transducer and trans activator STAT5A

Query Match 21.3%, Score 858, DB 2: Length 793;

Best Local Similarity 30.0%, Ident. NO. 3e-45;

Matches 245; Insertive 145; Mismatches 320; Indels 106; Gaps 28;

Oy 1 MOWNH--DTRRYKQIHLQYSDTFPMELRQFLAPWIESQDWAFA---ASKESHATIV 56

Db 1 MAGWV--GLQGDALRQMOVLYGQHPFIEVKRYTLAQWIESQPHNADIDNDQDKQATQL 60

Oy 57 FHNLI--IQQQYSRFLQESNVLYQHNLRRIKQFLOASRYLEKPMETIARTVARGCLWESRL 116

Db 61 LRGILWV--LQKKRKHQVQCEGDFILKIKICHYATQLONTYDRCPMELIVRCIRHLIVNEGRIV 120

Oy 117 QTAATP--VQVGGANPRTAAVYTERKQOMLQHLQVQRKKVQVQLEQKKMYENLQDDPDEN 176

Db 121 REANN--VGV---VLVDASQKHLQTNQRFEEIRLLITQDTFENLKLQDQTEFFIIOY 176

```

OY 177 -KTLKSQGMODLNGN-----QSVTRKMOOLEQML-----TALDOMRRSIVSELAC 224
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 177 QESLRIOAOFQALQOLNPOERMSRETLAQKQVSLFTWLQREAOITLQOYRVLEAEKHOKT 236

OY 225 LLSAMEYVQKTLTDEELADJMKRRRELACIGGPRNICLDRLENNITSLAESLOTROOKKL 284
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 225 LLSAMEYVQKTLTDEELADJMKRRRELACIGGPRNICLDRLENNITSLAESLOTROOKKL 284

OY 237 LQLLRKQOITLILDELOMRROOLAGNGPREGSLDVLSWCEKLAELTIWONROOIRRA 296
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 237 LQLLRKQOITLILDELOMRROOLAGNGPREGSLDVLSWCEKLAELTIWONROOIRRA 296

OY 285 EELQOKVSYKDPPIVOHRPMLERIVELFRNLKMSAFVVEROPCMHPRPLVIKTGYO 344
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 297 EHLQCOQLPIPG-PVEEMLAEVNATITDIIISALVSTFIEIKQ-----POLVKTQTK 347

OY 345 FTKRKVLVYKFPRELNYOL---KIKVCIDKDSGVAAALRSGRKN-----ILGTNTRKVMN 396
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 348 FAATVRLVVG-GRKLVNHNMPQVAKATIISEOAKSLKKNENTRNECSGEL--NCCVMEY 405

OY 397 EESNNGSLAEFKHLTREORCGNGRANCDAALVTEELHLITFETE--VYHOGKIDL 454
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 406 HQA-TGTLNHFNRMSLKRIRK-----RADRRGASVTEEKFTVLFEQSFVSGSNELVFOY 459

OY 455 ETHSLPVPVVISNICOMPANASILMYNMLTNKNVNFTRPGITGDQYAEVLISQFSS 514
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 460 KTLSPVPVIVHGSQDNNAATATVLMDNAFA-EPGRVP-FAVPDKVLPOLCEALNKKFKA 517

OY 515 -TTKRGSLTEQLTTLAELKLGPGVN-----YSGCQITMAKFCENMAGKGFSEFWVL 568
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 518 EVQSNRGLTKENLFLAOKLFPNISSNLELDYNSMSVMSQFNRENLPGMNTTFWQMFQGV 577

OY 569 IDLVKKTALANEGYIMGTISKERERAILSTKPPGTFLLRFSSSESGGVTFWVEKDI 628
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 578 MEVLKKHHRKHMNDGAILGFVNKQOAHDLINKPDGTFLLRFSD-SEIGGITIAW--KFD 634

OY 629 SGKTQIOQSVERTKQOOLNNMSFAEIIIMGYKIMDATNIIIVSPVLYLYDIPKEAEFGKYR 688
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 635 SPDRNLMLNLPFTTRDTSISLAD-----RLGD-----LNTLIVFPDRKDEVFAYTYT 684

OY 689 PE-----SQEHPE-----ADPSAARYLTKFKICVTPPTCSNTIDLPMSPT 730
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 685 PVLAKAVDGVKPOIKQVPEFVNASTDAGASATYMDQ---APSPVAVCPQP-HVNMYP 740

OY 731 LDSLMOFGNCGAEPASAGQFESLTFPMDLTSECA 766
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 741 PDPVPLD--QDGE-----FDLDESDMYA 760

RESULT 9
S55527
mammary gland factor - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 28-Jul-2000
C:Accession: S55527; S44353
R:Wakao, H.; Gouilleux, F.; Groner, B.
EMBO J. 14, 854-855, 1995
A:Title: Corrigenda. Mammmary gland factor (MGF) is a novel member of the cytokine regula
A:Reference number: S55527; MUID:95188889
A:Accession: S55527
A:Molecule type: mRNA
A:Residues: 1-794 <WAK>
A:Cross-references: EMBL:X78428; NID:9602354; PIDN:CA55191.1; PID:9602355
A>Note: this is a revision to the sequence from reference S44353
R:Wakao, H.; Gouilleux, F.; Groner, B.
EMBO J. 13, 2182-2191, 1994
A:Title: Mammmary gland factor (MGF) is a novel member of the cytokine regulated transcrip
A:Reference number: S44353; MUID:94244619
A:Accession: S44353
A:Molecule type: mRNA
A:Residues: 1-716, 'RHUGRGSLSR', 729, 'P', 731, 'ASL' <WAK>
A:Cross-references: EMBL:X78428
A>Note: this sequence has been revised in reference S55527
C:Superfamily: human signal transducer and transcription activator STAT5A

```

```

Query Match 19.68; Score 791.5; DB 2; Length 794;
Best Local Similarity 29.38; Pred. No. 3, 9e-41;
Matches 240; Conservative 146; Mismatches 323; Indels 109; Gaps 31;

OY 1 MAOMNLOQLDTRFYKLHOLYSDTFPMELOFLAPPIESODMAYA-----ASKESHATLY 56
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1 MACHTQAQLOGALQKMOVLYGQHFIEVRNHYLAQMIESQPMADALDIDNPRQAVOTL 60

OY 57 FNNLGEIDQYSRPIQESNVLYOHLN-RRIKQFLQSRVLEKMEIARIVAKCLMEESRL 115
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 61 LEGIVQELQKAEHQVCEDEFLKIKLGHVYVHSSRTTAAFPWSMLRCIRHILYNFORL 120

OY 116 LQTAATAQGGQGANPHTAAVTEKQOMLEQHLQDYKRVQDLEQKKVVENLQDDEDFN 175
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 121 VREKTNSSAG-----LVDAMQKHLQIQTTEELVYQDTEENELKKIQTOEFPIIO 176

OY 176 Y-KTLKSQGMODLNGN-----QSVTRKMOOLEQML-----TALDOMRRSIVSELAC 223
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 177 QESLRIOAOFQALQOLNPOERLSRETAQKQVSLFAMLQREAOITLQOYRVLEAEKHOK 236

OY 224 LLSAMEYVQKTLTDEELADJMKRRRELACIGGPRNICLDRLENNITSLAESLOTROOKKL 283
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 237 LQLLRKQOITLILDELOMRROOLAGNGPREGSLDVLSWCEKLAELTIWONROOIRR 295

OY 284 LLELOQVSYKDPPIVOHRPMLERIVELFRNLKMSAFVVEROPCMHPRPLVIKTGY 343
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 296 AEHLQCOQLPIPG-PVEEMLAEVNATITDIIISALVSTFIEIKQ-----POLVKTQTK 346

OY 344 QPTTKVNLVYKFPRELNYOL---KIKVCIDKDSGVAAALRSGRKN-----ILGTNTRKVMN 395
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 347 KFATVRLVVG-GRKLVNHNMPQVAKATIISEOAKSLKKNENTRNECSGEL--NCCVME 404

OY 396 MEESNNGSLAEFKHLTREORCGNGRANCDAALVTEELHLITFETE--VYHOGKIDL 453
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 405 YHQ-RTGTLNHFNRMSLKRIRK-----RADRRGASVTEEKFTVLFEQSFVSGSNELVFO 458

OY 454 LETHSLPVPVVISNICOMPANASILMYNMLTNKNVNFTRPGITGDQYAEVLISQFSS 513
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 459 VKTSLPVPVIVHGSQDNNAATATVLMDNAFA-EPGRVP-FAVPDKVLPOLCEALNKKF 516

OY 514 S--TTKRGSLTEQLTTLAELKLGPGVN-----YNSGCOITMAKFCENMAGKGFSEFWVL 567
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 517 AEVQSNRGLTKENLFLAOKLFPNSSSHLELDYNSMSVMSQFNRENLPGMNTTFWQMFQ 576

OY 568 IIDLVKKTALANEGYIMGTISKERERAILSTKPPGTFLLRFSSSESGGVTFWVEKDI 627
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 577 VMEVLKKHHRKHMNDGAILGFVNKQOAHDLINKPDGTFLLRFSD-SEIGGITIAW--KF 633

OY 628 ISGKTQIOQSVERTKQOOLNNMSFAEIIIMGYKIMDATNIIIVSPVLYLYDIPKEAEFGKY 687
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 634 DSPDRNLMLNLPFTTRDTSISLAD-----RLGD-----LNTLIVFPDRKDEVFSKY 683

OY 688 RPE-----SQEHPE-----AD-IGSAARYLTKFKICVTPPTCSNTIDLPMSPT 728
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 684 TPVLAKAVDGVKPOIKQVPEFVNASTDAGASATYMDQ---APSPVAVCPQP-HVNMYP 739

OY 729 RTDSLMOFGNCGAEPASAGQFESLTFPMDLTSECA 766
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 740 QNPDPVLD--QDGE-----FDLDETMYA 761

RESULT 10
I57557
DNA-Binding Protein and transcription factor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 28-Jul-2000
C:Accession: I57557
R:Quelle, F.W.; Shimoda, K.; Thieffelder, W.; Fischer, C.L.; Kim, A.; Ruben, S.M.; CI
Mol. Cell. Biol. 15, 3336-3343, 1995
A:Title: Cloning of murine Stat6, and human Stat6, Stat proteins that are tyrosine pho
A:Reference number: I57557; MUID:95280934
A:Accession: I57557
A>Status: preliminary; translated from GB/EMBL/DBJ

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 19, 2002, 14:30:33 ; Search time 23.01 Seconds
(without alignments)

753,044 Million cell updates/sec

Title: US-08-212-185-12

Perfect score: 4029
Sequence: 1 MAOMNOLQDLTRYIKQLHQ.....QFSLTFMDLTSCATSPM 770

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2.6/prodata/2/1aa/5A.COMB.pep.*
2: /cgn2.6/prodata/2/1aa/5B.COMB.pep.*
3: /cgn2.6/prodata/2/1aa/6A.COMB.pep.*
4: /cgn2.6/prodata/2/1aa/6B.COMB.pep.*
5: /cgn2.6/prodata/2/1aa/PCFUS.COMB.pep.*
6: /cgn2.6/prodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4029	100.0	770	1	US-08-369-796-12 Sequence 12, Appl
2	4029	100.0	770	2	US-08-852-091-12 Sequence 12, Appl
3	4029	100.0	770	3	US-08-820-754-12 Sequence 12, Appl
4	4029	100.0	770	3	US-08-956-652-12 Sequence 12, Appl
5	4029	100.0	770	3	US-08-956-869-12 Sequence 12, Appl
6	4029	100.0	770	3	US-08-948-547-12 Sequence 12, Appl
7	4029	100.0	770	4	US-09-364-970-3 Sequence 3, Appl
8	4029	100.0	770	5	PCT-US95-17025-12 Sequence 12, Appl
9	4018	99.7	770	3	US-09-012-710-8 Sequence 8, Appl
10	4016	99.7	770	4	US-09-364-970-5 Sequence 9, Appl
11	4006	99.4	770	1	US-08-416-5818-9 Sequence 5, Appl
12	3977	98.7	770	1	US-08-416-5818-1 Sequence 1, Appl
13	3977	98.7	770	1	US-08-416-5818-5 Sequence 5, Appl
14	3977	98.7	770	4	US-09-087-465-6 Sequence 6, Appl
15	3977	98.7	771	1	US-08-276-099A-14 Sequence 14, Appl
16	3977	98.7	771	1	US-08-781-890-14 Sequence 14, Appl
17	2017	50.1	750	1	US-08-369-796-4 Sequence 4, Appl
18	2017	50.1	750	2	US-08-852-091-4 Sequence 4, Appl
19	2017	50.1	750	2	US-08-820-754-4 Sequence 4, Appl
20	2017	50.1	750	3	US-08-956-652-4 Sequence 4, Appl
21	2017	50.1	750	3	US-08-956-869-4 Sequence 4, Appl
22	2017	50.1	750	3	US-08-948-547-4 Sequence 4, Appl
23	2017	50.1	750	3	US-09-087-465-2 Sequence 2, Appl
24	2017	50.1	750	4	US-09-364-970-1 Sequence 4, Appl
25	2017	50.1	750	5	PCT-US95-17025-4 Sequence 4, Appl
26	2004	49.7	750	4	US-09-364-970-8 Sequence 8, Appl
27	1997	49.6	712	1	US-08-369-796-6 Sequence 6, Appl

28	1997	49.6	712	2	US-08-852-091-6	Sequence 6, Appl
29	1997	49.6	712	2	US-08-820-754-6	Sequence 6, Appl
30	1997	49.6	712	3	US-08-956-652-6	Sequence 6, Appl
31	1997	49.6	712	3	US-08-956-869-6	Sequence 6, Appl
32	1997	49.6	712	3	US-08-948-547-6	Sequence 6, Appl
33	1997	49.6	712	5	PCT-US95-17025-6	Sequence 6, Appl
34	1984	49.2	749	1	US-08-369-796-8	Sequence 8, Appl
35	1984	49.2	749	2	US-08-852-091-8	Sequence 8, Appl
36	1984	49.2	749	3	US-08-820-754-8	Sequence 8, Appl
37	1984	49.2	749	3	US-08-956-652-8	Sequence 8, Appl
38	1984	49.2	749	3	US-08-956-869-8	Sequence 8, Appl
39	1984	49.2	749	3	US-09-012-710-7	Sequence 7, Appl
40	1984	49.2	749	3	US-08-948-547-8	Sequence 8, Appl
41	1984	49.2	749	5	PCT-US95-17025-8	Sequence 8, Appl
42	1971.5	48.9	740	1	US-08-276-099A-12	Sequence 12, Appl
43	1971.5	48.9	740	1	US-08-781-890-12	Sequence 12, Appl
44	1798	44.6	748	1	US-08-408-318-2	Sequence 2, Appl
45	1798	44.6	748	1	US-08-839-164-2	Sequence 2, Appl

ALIGNMENTS

```
RESULT 1
US-08-369-796-12
: Sequence 12, Application US/08369796
: Patent No. 5716622
:
: GENERAL INFORMATION:
: APPLICANT: James E. Darnell, Jr.
: APPLICANT: Zilong Wen
: APPLICANT: Curt M. Horvath
: APPLICANT: Zhong Zhong
: TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
: TITLE OF INVENTION: TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
: NUMBER OF SEQUENCES: 39
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Klauber & Jackson
: STREET: 411 Hackensack Avenue
: CITY: Hackensack
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07601
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.0
:
: CURRENT PATENT DATA:
: APPLIC NUMBER: US/08/369,796
: FILING NUMBER: 06-JAN-1995
: CLASSIFICATION: 4: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Jackson Esq., David A.
: REGISTRATION NUMBER: 26,742
: REFERENCE CITED INFORMATION:
: TELECOMMUNICATION NUMBER: 600-1-116
: TELEPHONE: 201 487-5800
: TELEFAX: 201 343-1684
:
: INFORMATION: SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 770 amino acids
: TYPE: coding acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
: US-08-369-796-12
```

Query Match 100.0%; Score 4029; DB 1; Length 770;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Caps 0;

0y 1 MAOMNOLQDLTRYIKQLHQ..YSDTFPMLRQFLAPWIESODMAYASKESHATLVFNL 60

```
|||||
Db 1 MAOMNQLODLTRYKQLHQLYSDFPMEILKQFLAPWIESQDMVAASKESHATLVFPHNL 60
QY 61 LGEIDDOYSRFLQESNVLYOHNLRRIKOFLQSRYLEKPMETARIYARCLWESRLQTA 120
Db 61 LGEIDDOYSRFLQESNVLYOHNLRRIKOFLQSRYLEKPMETARIYARCLWESRLQTA 120
QY 121 TAAOOGGQANHPTAAVVEKEQOMLEQHLQDYKRVODLEOKMKVVENLQDDFDFNYKTLK 180
Db 121 TAAOOGGQANHPTAAVVEKEQOMLEQHLQDYKRVODLEOKMKVVENLQDDFDFNYKTLK 180
QY 181 SOGDMQDLNGNNSVTRQKMOQLEQMLTALDQMRRSIVSELAGLSAMEYVOKTLTDEEL 240
Db 181 SOGDMQDLNGNNSVTRQKMOQLEQMLTALDQMRRSIVSELAGLSAMEYVOKTLTDEEL 240
QY 241 ADMKRRPEIACIGCPNCLDRLENNITSLAESQLQTRQOIKKELELOOKVSYKGDPIVQ 300
Db 241 ADMKRRPEIACIGCPNCLDRLENNITSLAESQLQTRQOIKKELELOOKVSYKGDPIVQ 300
QY 301 HRPMLEERIVELFRNLMSAFYVERQPCMPHDPRLVKTGVQFTTKVRLVFPPELNY 360
Db 301 HRPMLEERIVELFRNLMSAFYVERQPCMPHDPRLVKTGVQFTTKVRLVFPPELNY 360
QY 361 QLKIKVCIKDSGDVAALRGSRKFNILGNTKVMNMEESNNGSLSAEFKHLTLREORCGN 420
Db 361 QLKIKVCIKDSGDVAALRGSRKFNILGNTKVMNMEESNNGSLSAEFKHLTLREORCGN 420
QY 421 GGRACDASLIYTEELHLITFEFEYHOGCLKIDLETHSLPVVVISNICOMPANASILMY 480
Db 421 GGRACDASLIYTEELHLITFEFEYHOGCLKIDLETHSLPVVVISNICOMPANASILMY 480
QY 481 NMLTNPNKVNNEFTKPPIGTMDOVAEVLWSQFSSTTKRGLSIEQLTTLAEKLLGPGVNS 540
Db 481 NMLTNPNKVNNEFTKPPIGTMDOVAEVLWSQFSSTTKRGLSIEQLTTLAEKLLGPGVNS 540
QY 541 GCQITWAFCKENMAGKGFSPVWMLDNIIDLKYYILALMNGYIMGFIISKEREALIST 600
Db 541 GCQITWAFCKENMAGKGFSPVWMLDNIIDLKYYILALMNGYIMGFIISKEREALIST 600
QY 601 KPPGTFLLRFSESSKEGVTFTWVEKDLSGKTQIOSVEPYTKQOLNNNSFAEIIINGYKIM 660
Db 601 KPPGTFLLRFSESSKEGVTFTWVEKDLSGKTQIOSVEPYTKQOLNNNSFAEIIINGYKIM 660
QY 661 DATNIVLSPLYLYPDIPKEAFKCYCRPESOEHPDGPAAPIKTKFCVTPPTCSN 720
Db 661 DATNIVLSPLYLYPDIPKEAFKCYCRPESOEHPDGPAAPIKTKFCVTPPTCSN 720
QY 721 TIDLPMSPRTLDLSLMQFNGNGGAPPSAGGOFESLTFPMDLTSECATSPM 770
Db 721 TIDLPMSPRTLDLSLMQFNGNGGAPPSAGGOFESLTFPMDLTSECATSPM 770

RESULT 2
US-08-852-091-12
Sequence 12 Application US/08852091
Patent No. 5683228
GENERAL INFORMATION:
APPLICANT: James E. Darnell, Jr.
APPLICANT: Zilong Wen
APPLICANT: Curt M. Horvath
APPLICANT: Zhong Zhong
TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESS: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: Hackensack
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
```

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852,091
FILING DATE: 06-MAY-1997
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/369,796
FILING DATE: 06-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-116
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 770 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-852-091-12
```

```
Query Match 100.0% Score 4029 DB 2 Length 770;
Best Local Similarity 100.0% Pred. No. 0;
Matches 770: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 MAOMNQLODLTRYKQLHQLYSDFPMEILKQFLAPWIESQDMVAASKESHATLVFPHNL 60
Db 1 MAOMNQLODLTRYKQLHQLYSDFPMEILKQFLAPWIESQDMVAASKESHATLVFPHNL 60
QY 61 LGEIDDOYSRFLQESNVLYOHNLRRIKOFLQSRYLEKPMETARIYARCLWESRLQTA 120
Db 61 LGEIDDOYSRFLQESNVLYOHNLRRIKOFLQSRYLEKPMETARIYARCLWESRLQTA 120
QY 121 TAAOOGGQANHPTAAVVEKEQOMLEQHLQDYKRVODLEOKMKVVENLQDDFDFNYKTLK 180
Db 121 TAAOOGGQANHPTAAVVEKEQOMLEQHLQDYKRVODLEOKMKVVENLQDDFDFNYKTLK 180
QY 181 SOGDMQDLNGNNSVTRQKMOQLEQMLTALDQMRRSIVSELAGLSAMEYVOKTLTDEEL 240
Db 181 SOGDMQDLNGNNSVTRQKMOQLEQMLTALDQMRRSIVSELAGLSAMEYVOKTLTDEEL 240
QY 241 ADMKRRPEIACIGCPNCLDRLENNITSLAESQLQTRQOIKKELELOOKVSYKGDPIVQ 300
Db 241 ADMKRRPEIACIGCPNCLDRLENNITSLAESQLQTRQOIKKELELOOKVSYKGDPIVQ 300
QY 301 HRPMLEERIVELFRNLMSAFYVERQPCMPHDPRLVKTGVQFTTKVRLVFPPELNY 360
Db 301 HRPMLEERIVELFRNLMSAFYVERQPCMPHDPRLVKTGVQFTTKVRLVFPPELNY 360
QY 361 QLKIKVCIKDSGDVAALRGSRKFNILGNTKVMNMEESNNGSLSAEFKHLTLREORCGN 420
Db 361 QLKIKVCIKDSGDVAALRGSRKFNILGNTKVMNMEESNNGSLSAEFKHLTLREORCGN 420
QY 421 GGRACDASLIYTEELHLITFEFEYHOGCLKIDLETHSLPVVVISNICOMPANASILMY 480
Db 421 GGRACDASLIYTEELHLITFEFEYHOGCLKIDLETHSLPVVVISNICOMPANASILMY 480
QY 481 NMLTNPNKVNNEFTKPPIGTMDOVAEVLWSQFSSTTKRGLSIEQLTTLAEKLLGPGVNS 540
Db 481 NMLTNPNKVNNEFTKPPIGTMDOVAEVLWSQFSSTTKRGLSIEQLTTLAEKLLGPGVNS 540
QY 541 GCQITWAFCKENMAGKGFSPVWMLDNIIDLKYYILALMNGYIMGFIISKEREALIST 600
Db 541 GCQITWAFCKENMAGKGFSPVWMLDNIIDLKYYILALMNGYIMGFIISKEREALIST 600
QY 601 KPPGTFLLRFSESSKEGVTFTWVEKDLSGKTQIOSVEPYTKQOLNNNSFAEIIINGYKIM 660
Db 601 KPPGTFLLRFSESSKEGVTFTWVEKDLSGKTQIOSVEPYTKQOLNNNSFAEIIINGYKIM 660
```

DB 601 KPPGFTLLRFSSSSKEGCVFTTWEEKDISGKTQIOSVEPYTKOOLNMSFAELIMGYKM 660
QY 661 DATNIIIVSLVLYLPDIPEKEAFGKYCRPESQEHPEADGSAAPYLKTKFCVPTTCSN 720
DB 661 DATNIIIVSLVLYLPDIPEKEAFGKYCRPESQEHPEADGSAAPYLKTKFCVPTTCSN 720
QY 721 TIDLPMSPTLDSLMQFGNNGEAGPSAGGPFESLTFDMDLTSECATSPM 770
DB 721 TIDLPMSPTLDSLMQFGNNGEAGPSAGGPFESLTFDMDLTSECATSPM 770

RESULT 3
US-08-820-754-12
Sequence 12, Application US/08820754
Patent No. 5978835
GENERAL INFORMATION:
APPLICANT: Darnell Jr., James E.
APPLICANT: Schindler, Christian W.
APPLICANT: Fu, Xian-Yuan
APPLICANT: Wen, Zilong
APPLICANT: Zhong, Zhong
TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/820,754
FILING DATE: 19-MAR-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,185
FILING DATE: 11-MAR-1994
APPLICATION NUMBER: US 07/980,498
FILING DATE: 23-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/854,296
FILING DATE: 19-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO US93/02569
FILING DATE: 19-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,588
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-073 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 770 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-820-754-12

Query Match 100.0%; Score 4029; DB 2: Length 770;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 770: insertive 0: mismatches 0: indels 0: gaps 0:
QY 1 MAQMNLOOLDIIRYLYKOLHOLYSDFPMELOPLAPWIESQOMAYAAKESHATLVPHNL 60
DB 1 MAQMNLOOLDIIRYLYKOLHOLYSDFPMELOPLAPWIESQOMAYAAKESHATLVPHNL 60
QY 61 LGEIDOOVSRLFQESNVLYOHNLRRIKOFLORSYLEKPEIARIRARCIWESRLQTA 120
DB 61 LGEIDOOVSRLFQESNVLYOHNLRRIKOFLORSYLEKPEIARIRARCIWESRLQTA 120
QY 121 TAAOQGGQANHPPTAAVVEFEKQOMLEOHLODVARRVODLKOKKKVVENLODDDFNFKTLK 180
DB 121 TAAOQGGQANHPPTAAVVEFEKQOMLEOHLODVARRVODLEQKKVVENLODDDFNFKTLK 180
QY 181 SOGDMDLGNNOVSVRQAMOOLEOMLTALDOMKRSIVSELAGLISAMEYVOKTLTDEEL 240
DB 181 SOGDMDLGNNOVSVRQAMOOLEOMLTALDOMKRSIVSELAGLISAMEYVOKTLTDEEL 240
QY 241 ADMKRRPEIACIGPPNICLDRLENNITSLAESOLOTROQIKKLELOQKVSXKGDPIVO 300
DB 241 ADMKRRPEIACIGPPNICLDRLENNITSLAESOLOTROQIKKLELOQKVSXKGDPIVO 300
QY 301 HRPMLFERIVLEFRNIMKSAFYVEROPCPMHHPDRPLVAKTSVOFTTKVRLLVKPEPLNY 360
DB 301 HRPMLFERIVLEFRNIMKSAFYVEROPCPMHHPDRPLVAKTSVOFTTKVRLLVKPEPLNY 360
QY 361 OLKIKVCIDKSDGVVALRGSRKFNLCNTNKMNMESNNOSLSAFPHLTLRORCGN 420
DB 361 OLKIKVCIDKSDGVVALRGSRKFNLCNTNKMNMESNNOSLSAFPHLTLRORCGN 420
QY 421 GGRANDASLIVTEELHLITFETEVYHOGIKIDLETHISIPVVVISNICOMPAMASITLWY 480
DB 421 GGRANDASLIVTEELHLITFETEVYHOGIKIDLETHISIPVVVISNICOMPAMASITLWY 480
QY 481 NMILTNPKNVNPFTKPPITGTDQVAEIVSMQPSSTTKRGLSTHOUTTAEKILGCVNVS 540
DB 481 NMILTNPKNVNPFTKPPITGTDQVAEIVSMQPSSTTKRGLSTHOUTTAEKILGCVNVS 540
QY 541 GCOITWAKFKENMAGKGSFWMLDNIIDLVKKYIIALMNGVIMGFSKREBAIIST 600
DB 541 GCOITWAKFKENMAGKGSFWMLDNIIDLVKKYIIALMNGVIMGFSKREBAIIST 600
QY 601 KPPGFTLLRFSSSSKEGCVFTTWEEKDISGKTQIOSVEPYTKOOLNMSFAELIMGYKM 660
DB 601 KPPGFTLLRFSSSSKEGCVFTTWEEKDISGKTQIOSVEPYTKOOLNMSFAELIMGYKM 660
QY 661 DATNIIIVSLVLYLPDIPEKEAFGKYCRPESQEHPEADGSAAPYLKTKFCVPTTCSN 720
DB 661 DATNIIIVSLVLYLPDIPEKEAFGKYCRPESQEHPEADGSAAPYLKTKFCVPTTCSN 720
QY 721 TIDLPMSPTLDSLMQFGNNGEAGPSAGGPFESLTFDMDLTSECATSPM 770
DB 721 TIDLPMSPTLDSLMQFGNNGEAGPSAGGPFESLTFDMDLTSECATSPM 770

RESULT 4
US-08-956-652-12
Sequence 12, Application US/08956652
Patent No. 6013475
GENERAL INFORMATION:
APPLICANT: Darnell Jr., James E.
APPLICANT: Schindler, Christian W.
APPLICANT: Fu, Xian-Yuan
APPLICANT: Wen, Zilong
APPLICANT: Zhong, Zhong
TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey

```

: COUNTRY: USA
: ZIP: 07601
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/956,652
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/212,185
: FILING DATE: 11-MAR-1994
: APPLICATION NUMBER: US 07/980,498
: FILING DATE: 23-NOV-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/854,296
: FILING DATE: 19-MAR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO US93/02569
: FILING DATE: 19-MAR-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/126,588
: FILING DATE: 24-SEP-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Jackson Esq., David A.
: REGISTRATION NUMBER: 26,742
: REFERENCE/DOCKET NUMBER: 600-1-073 CIP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201 487-5800
: TELEFAX: 201 343-1684
: TELEX: 133521
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 770 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-956-652-12

```

Query Match 100.0%; Score 4029; DB 3; length 770;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MAOMNLOOLDTRYLKQHLQYSDTFPMELROFLAPWIESODMAAASKESMTLVFPHNL 60
DB 1 MAOMNLOOLDTRYLKQHLQYSDTFPMELROFLAPWIESODMAAASKESMTLVFPHNL 60
OY 61 LGEIDQYRSFLOESNVLVQHNLRRIKQFLOSRYLEKPMELIARIVARCLMEESRLQTA 120
DB 61 LGEIDQYRSFLOESNVLVQHNLRRIKQFLOSRYLEKPMELIARIVARCLMEESRLQTA 120
OY 121 TAAQGGGQANHPTAAVYTEKQMLECHLQDVRRVODLEQKMKVVENLQDDPFNTKTLK 180
DB 121 TAAQGGGQANHPTAAVYTEKQMLECHLQDVRRVODLEQKMKVVENLQDDPFNTKTLK 180
OY 181 SOGDMDLNGNNSVTRQKMOOLEOMLTALDOMRSIVSLAGLSAMEYVOKTLDEEL 240
DB 181 SOGDMDLNGNNSVTRQKMOOLEOMLTALDOMRSIVSLAGLSAMEYVOKTLDEEL 240
OY 241 ADMKRRPEIACIGPPNICDLRLNMTTSLAESOLOTRQOIKKLEELQOKVSYKGPYVQ 300
DB 241 ADMKRRPEIACIGPPNICDLRLNMTTSLAESOLOTRQOIKKLEELQOKVSYKGPYVQ 300
OY 301 HRPMLEERIVLEFRNLKMSAFVYERQPCMHDRPLVITGVQFTTKVALLKFPPLNY 360
DB 301 HRPMLEERIVLEFRNLKMSAFVYERQPCMHDRPLVITGVQFTTKVALLKFPPLNY 360
OY 361 QLKIKVCIDKSDGVALLRSRKFNIIGTNTKYNMESNNGSLSAEFKHLTLREORCGN 420
DB 361 QLKIKVCIDKSDGVALLRSRKFNIIGTNTKYNMESNNGSLSAEFKHLTLREORCGN 420

```

```

OY 421 GGRANDASLIYTELHLITFETEVYHOGKIDLETHSLPVVVISNCOMPMANASILMY 480
DB 421 GGRANDASLIYTELHLITFETEVYHOGKIDLETHSLPVVVISNCOMPMANASILMY 480
OY 481 NMLTNNPKVNFETKPPIGTWQVAEVLWSOPSSYTKRGLSTEOLTTAEKLLGCVNYS 540
DB 481 NMLTNNPKVNFETKPPIGTWQVAEVLWSOPSSYTKRGLSTEOLTTAEKLLGCVNYS 540
OY 541 GCQITWAKFCKENMAGKGFSPWMLDNIIDLVKKYILALMNEGYIMGFLSKRERAILST 600
DB 541 GCQITWAKFCKENMAGKGFSPWMLDNIIDLVKKYILALMNEGYIMGFLSKRERAILST 600
OY 601 KPPGFTLLRFSSSKEGVTFTWVEKDISGKTQIOSVEPYTQOOLNNSFAELIMGYKIM 660
DB 601 KPPGFTLLRFSSSKEGVTFTWVEKDISGKTQIOSVEPYTQOOLNNSFAELIMGYKIM 660
OY 661 DATNILVSPVLYLPDIPKEAFKCYCRPESQHPHADGSAAPYLKTKFICVPTTCSN 720
DB 661 DATNILVSPVLYLPDIPKEAFKCYCRPESQHPHADGSAAPYLKTKFICVPTTCSN 720
OY 721 TIDLPMSPRTLDLSMOFGNNGGAEPSSAGOPRESLTFMDLTSECATSPM 770
DB 721 TIDLPMSPRTLDLSMOFGNNGGAEPSSAGOPRESLTFMDLTSECATSPM 770

```

RESULT 5
 US-08-956-869-12
 : Sequence 12, Application US/08956869
 : Patent No. 6030808
 : GENERAL INFORMATION:
 : APPLICANT: Darnell Jr., James E.
 : APPLICANT: Schindler, Christian W.
 : APPLICANT: Fu, Xian-Yuan
 : APPLICANT: Wen, Zilong
 : TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
 : TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF
 : NUMBER OF SEQUENCES: 25
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Klauber & Jackson
 : STREET: 411 Hackensack Avenue
 : CITY: Hackensack
 : STATE: New Jersey
 : COUNTRY: USA
 : ZIP: 07601
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/956,869
 : FILING DATE:
 : CLASSIFICATION:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/212,185
 : FILING DATE:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/854,296
 : FILING DATE: 19-MAR-1992
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: WO US93/02569
 : FILING DATE: 19-MAR-1993
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/126,588
 : FILING DATE: 24-SEP-1993
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Jackson Esq., David A.
 : REGISTRATION/DOCKET NUMBER: 26,742
 : REFERENCE/DOCKET NUMBER: 600-1-073 CIP
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684
 TELEX: 133521
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 770 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-956-869-12

Query Match 100.0%; Score 4029; DB 3; Length 770;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MAONNOLOOUDTRYKOLHOLYSOTFPMELROFLAPWIESODMAYAAKESHATLVFNNL 60
 DB 1 MAONNOLOOUDTRYKOLHOLYSOTFPMELROFLAPWIESODMAYAAKESHATLVFNNL 60
 OY 61 LGEIDOOVSRLFOESNVLYOHNLRRIKOFLORSYLEKPMELIARIYARCLMEESRLLOTA 120
 DB 61 LGEIDOOVSRLFOESNVLYOHNLRRIKOFLORSYLEKPMELIARIYARCLMEESRLLOTA 120
 OY 121 TAAOOGGGAHPTAAVYTEKQOMLEOHLODYRRKRYODLEOKMKVVENLODDDFNYKTLLK 180
 DB 121 TAAOOGGGAHPTAAVYTEKQOMLEOHLODYRRKRYODLEOKMKVVENLODDDFNYKTLLK 180
 OY 181 SOGDMODLNGNOSVTRQKMOOLEOMLTALDOMRSTYSELAGLLSMEYOKTLTDEEL 240
 DB 181 SOGDMODLNGNOSVTRQKMOOLEOMLTALDOMRSTYSELAGLLSMEYOKTLTDEEL 240
 OY 241 ADMKRREPIACIGPPNICDRLENNITSLAESOLQTRQOKKLEELDOOKSYKCDPIYQ 300
 DB 241 ADMKRREPIACIGPPNICDRLENNITSLAESOLQTRQOKKLEELDOOKSYKCDPIYQ 300
 OY 301 HRPMLEERIVELFRNLKMSAFVEROPCPMHPDRPLVKTGOVFTTKVRLVKEPELNY 360
 DB 301 HRPMLEERIVELFRNLKMSAFVEROPCPMHPDRPLVKTGOVFTTKVRLVKEPELNY 360
 OY 361 OLKIKVCIDKSGVAAALRGSRKFENILGTNTKVMMEESNNGSLSAEKKHLTLREORGN 420
 DB 361 OLKIKVCIDKSGVAAALRGSRKFENILGTNTKVMMEESNNGSLSAEKKHLTLREORGN 420
 OY 421 GGRANCASLIVTEELHLITFEFEVYHOGKLTIDETHSLPVVVISNICOMPMAASILMY 480
 DB 421 GGRANCASLIVTEELHLITFEFEVYHOGKLTIDETHSLPVVVISNICOMPMAASILMY 480
 OY 481 NMLTNNRNKNVFTKPPIGTDQVAEVLSPQSSSTTKRGCLSTEOLITLAEKLLGCVNYS 540
 DB 481 NMLTNNRNKNVFTKPPIGTDQVAEVLSPQSSSTTKRGCLSTEOLITLAEKLLGCVNYS 540
 OY 541 GCOTITMAKFCCKENNAKGFSPWVWLDNITDLYKKYLLALMEGYMGFTSKEREALIST 600
 DB 541 GCOTITMAKFCCKENNAKGFSPWVWLDNITDLYKKYLLALMEGYMGFTSKEREALIST 600
 OY 601 KPPGTFLLRFSESSKEGCVTFETWEKDISGKTQIOSVEPYTKQOLNNNSFAEIIIMGYKIM 660
 DB 601 KPPGTFLLRFSESSKEGCVTFETWEKDISGKTQIOSVEPYTKQOLNNNSFAEIIIMGYKIM 660
 OY 661 DATNIVLSPLVYLPDIKPEAFKCYKCRPESEOEHPADPGSAAPLKKKFCVPTTCSN 720
 DB 661 DATNIVLSPLVYLPDIKPEAFKCYKCRPESEOEHPADPGSAAPLKKKFCVPTTCSN 720
 OY 721 TIDLPMSRITLDSLMQFCNGNEGAEPASAGOFESTLFDMDLTSECATSPM 770
 DB 721 TIDLPMSRITLDSLMQFCNGNEGAEPASAGOFESTLFDMDLTSECATSPM 770

RESULT 6
 US-08-948-547-12
 Sequence 12, Application us/08948547
 Patent No. 6124118
 GENERAL INFORMATION:

APPLICANT: Arnell Jr., James E.
 APPLICANT: Schindler, Christian W.
 APPLICANT: Fu, Xian-Yuan
 APPLICANT: Men, Zilong
 APPLICANT: Zhong, Zhong
 TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
 NUMBER OF SEQUENCES: 25
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Klauber & Jackson
 STREET: 411 Hackensack Avenue
 CITY: Hackensack
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07601
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/948,547
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/212,185
 FILING DATE: 11-MAR-1994
 APPLICATION NUMBER: US/07/980,498
 FILING DATE: 23-NOV-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/854,296
 FILING DATE: 19-MAR-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO/US93/02569
 FILING DATE: 19-MAR-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/126,588
 FILING DATE: 24-SEP-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Jackson Esq., David A.
 REGISTRATION NUMBER: 26,742
 REFERENCE/MARKET NUMBER: 600-1-073 CIP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201 487-5800
 TELEFAX: 201 343-1684
 TELEX: 133521
 INFORMATION: SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 770 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-948-547-12

Query Match 100.0%; Score 4029; DB 3; Length 770;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MAONNOLOOUDTRYKOLHOLYSOTFPMELROFLAPWIESODMAYAAKESHATLVFNNL 60
 DB 1 MAONNOLOOUDTRYKOLHOLYSOTFPMELROFLAPWIESODMAYAAKESHATLVFNNL 60
 OY 61 LGEIDOOVSRLFOESNVLYOHNLRRIKOFLORSYLEKPMELIARIYARCLMEESRLLOTA 120
 DB 61 LGEIDOOVSRLFOESNVLYOHNLRRIKOFLORSYLEKPMELIARIYARCLMEESRLLOTA 120
 OY 121 TAAOOGGGAHPTAAVYTEKQOMLEOHLODYRRKRYODLEOKMKVVENLODDDFNYKTLLK 180
 DB 121 TAAOOGGGAHPTAAVYTEKQOMLEOHLODYRRKRYODLEOKMKVVENLODDDFNYKTLLK 180
 OY 181 SOGDMODLNGNOSVTRQKMOOLEOMLTALDOMRSTYSELAGLLSMEYOKTLTDEEL 240
 DB 181 SOGDMODLNGNOSVTRQKMOOLEOMLTALDOMRSTYSELAGLLSMEYOKTLTDEEL 240

Db 181 SOGDMODLNGNNSVTRQKMOOLEOMLTALDOMRRSIVSELAGLLSAMRYOKLTUDEEL 240
QY 241 ADMKRRPEIACIGPPNICDLRLENMTISLAESOLQTRQIITKLEELQOKYSKGPDIYQ 300
Db 241 ADMKRRPEIACIGPPNICDLRLENMTISLAESOLQTRQIITKLEELQOKYSKGPDIYQ 300
QY 301 HRPMLEERIVELFRNLKMSAFVEROPCPMHPDRPLVIKTGVQFTTKVRLVLPPELNY 360
Db 301 HRPMLEERIVELFRNLKMSAFVEROPCPMHPDRPLVIKTGVQFTTKVRLVLPPELNY 360
QY 361 OLKIKVCIDKDSGDVAALRGSRKFENILGTNTKVMNMESSNNGSLSAEFKHLLTREORCGN 420
Db 361 OLKIKVCIDKDSGDVAALRGSRKFENILGTNTKVMNMESSNNGSLSAEFKHLLTREORCGN 420
QY 421 GGRANDASLIYTEBLHITFETEVYHOGKIDLETHSLPVVVISNICOMPNAASILMY 480
Db 421 GGRANDASLIYTEBLHITFETEVYHOGKIDLETHSLPVVVISNICOMPNAASILMY 480
QY 481 NMLTNPKNVNPFTRKPIGTWDOVAEVLSPQSSSTTKRGLSIEOLITLAEKLLGPGVNS 540
Db 481 NMLTNPKNVNPFTRKPIGTWDOVAEVLSPQSSSTTKRGLSIEOLITLAEKLLGPGVNS 540
QY 541 GCQITWAFCKENMAGKGFSEFWMLDNIIDLKRYILALMNEGYIMGFIISKERERAILST 600
Db 541 GCQITWAFCKENMAGKGFSEFWMLDNIIDLKRYILALMNEGYIMGFIISKERERAILST 600
QY 601 KPPGFTLLRFSESSKEGCVTFWTEKDISGKTQIOSVEPYTKQOLNNNSFAEIIIMGYKIM 660
Db 601 KPPGFTLLRFSESSKEGCVTFWTEKDISGKTQIOSVEPYTKQOLNNNSFAEIIIMGYKIM 660
QY 661 DATNLVSPLYLVLPDIPKEBAFGKYCRPESOEHPDPSAAPYLKTKFCVPTTCSN 720
Db 661 DATNLVSPLYLVLPDIPKEBAFGKYCRPESOEHPDPSAAPYLKTKFCVPTTCSN 720
QY 721 TIDLPMSPRTLDSLMQFNGNNEGAPPSAGGFESLTFDMDLTSECATSPM 770
Db 721 TIDLPMSPRTLDSLMQFNGNNEGAPPSAGGFESLTFDMDLTSECATSPM 770

RESULT 7
US-09-364-970-3
: Sequence 3, Application US/09364970
: Patent No. 6235873
: GENERAL INFORMATION:
: APPLICANT: Bromberg, Jacqueline
: TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR
: TITLE OF INVENTION: DYSPLASIA-INDUCING MODULATORS OF ACTIVITY INCLUDING
: FILE REFERENCE: 600-1-252
: CURRENT APPLICATION NUMBER: US/09/364,970
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: Patent In Ver. 2.0
: SEQ ID NO 3
: LENGTH: 770
: TYPE: PRT
: ORGANISM: Mus musculus
US-09-364-970-3

Query Match 100.0% Score 4029; DB 4: Length 770;
Best Local Similarity 100.0% Pred. No. 0;
Matches 770: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAOMNLOOLDLTRYKOLHOLYSDTFPEMLROFLAPMIESODMAYAAKESHATLVFNL 60
Db 1 MAOMNLOOLDLTRYKOLHOLYSDTFPEMLROFLAPMIESODMAYAAKESHATLVFNL 60
QY 61 LGEIDQOYSRFLQESNVLYOHNLRLRIKQFLQSRYLEKPEMEIARIVARCLMEESRLQTA 120
Db 61 LGEIDQOYSRFLQESNVLYOHNLRLRIKQFLQSRYLEKPEMEIARIVARCLMEESRLQTA 120
QY 121 TAAOOGGQANHPRTAAVYVTEKQOMLEOHLODVAKKRVODLEQKKAVVENLODDPFPNTLK 180

Db 121 TAAOOGGQANHPRTAAVYVTEKQOMLEOHLODVAKKRVODLEQKKAVVENLODDPFPNTLK 180
QY 181 SOGDMODLNGNNSVTRQKMOOLEOMLTALDOMRRSIVSELAGLLSAMRYOKLTUDEEL 240
Db 181 SOGDMODLNGNNSVTRQKMOOLEOMLTALDOMRRSIVSELAGLLSAMRYOKLTUDEEL 240
QY 241 ADMKRRPEIACIGPPNICDLRLENMTISLAESOLQTRQIITKLEELQOKYSKGPDIYQ 300
Db 241 ADMKRRPEIACIGPPNICDLRLENMTISLAESOLQTRQIITKLEELQOKYSKGPDIYQ 300
QY 301 HRPMLEERIVELFRNLKMSAFVEROPCPMHPDRPLVIKTGVQFTTKVRLVLPPELNY 360
Db 301 HRPMLEERIVELFRNLKMSAFVEROPCPMHPDRPLVIKTGVQFTTKVRLVLPPELNY 360
QY 361 OLKIKVCIDKDSGDVAALRGSRKFENILGTNTKVMNMESSNNGSLSAEFKHLLTREORCGN 420
Db 361 OLKIKVCIDKDSGDVAALRGSRKFENILGTNTKVMNMESSNNGSLSAEFKHLLTREORCGN 420
QY 421 GGRANDASLIYTEBLHITFETEVYHOGKIDLETHSLPVVVISNICOMPNAASILMY 480
Db 421 GGRANDASLIYTEBLHITFETEVYHOGKIDLETHSLPVVVISNICOMPNAASILMY 480
QY 481 NMLTNPKNVNPFTRKPIGTWDOVAEVLSPQSSSTTKRGLSIEOLITLAEKLLGPGVNS 540
Db 481 NMLTNPKNVNPFTRKPIGTWDOVAEVLSPQSSSTTKRGLSIEOLITLAEKLLGPGVNS 540
QY 541 GCQITWAFCKENMAGKGFSEFWMLDNIIDLKRYILALMNEGYIMGFIISKERERAILST 600
Db 541 GCQITWAFCKENMAGKGFSEFWMLDNIIDLKRYILALMNEGYIMGFIISKERERAILST 600
QY 601 KPPGFTLLRFSESSKEGCVTFWTEKDISGKTQIOSVEPYTKQOLNNNSFAEIIIMGYKIM 660
Db 601 KPPGFTLLRFSESSKEGCVTFWTEKDISGKTQIOSVEPYTKQOLNNNSFAEIIIMGYKIM 660
QY 661 DATNLVSPLYLVLPDIPKEBAFGKYCRPESOEHPDPSAAPYLKTKFCVPTTCSN 720
Db 661 DATNLVSPLYLVLPDIPKEBAFGKYCRPESOEHPDPSAAPYLKTKFCVPTTCSN 720
QY 721 TIDLPMSPRTLDSLMQFNGNNEGAPPSAGGFESLTFDMDLTSECATSPM 770
Db 721 TIDLPMSPRTLDSLMQFNGNNEGAPPSAGGFESLTFDMDLTSECATSPM 770

RESULT 8
PCT-US95-17025-12
: Sequence 12, Application PC/TUS9517025
: GENERAL INFORMATION:
: APPLICANT: James E. Darnell, Jr.
: APPLICANT: Zilong Wen
: APPLICANT: Curt M. Horvath
: APPLICANT: Zhong Zhong
: TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
: TITLE OF INVENTION: TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
: NUMBER OF SEQUENCES: 39
: CORRESPONDENT ADDRESS:
: ADDRESS: Klauder & Jackson
: STREET: 401 Hackensack Avenue
: CITY: Hackensack
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07601
: COMPUTER REPRODUCIBLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/17025
: FILING DATE: 28-DEC-1995
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/369,796

```

: FILING DATE: 06-JAN-1995
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Jackson Esq., David A.
: REGISTRATION NUMBER: 26,742
: REFERENCE/DOCKET NUMBER: 600-1-116
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201 487-5800
: TELEFAX: 201 343-1684
: TELEX: 133521
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 770 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PCT-US95-17025-12

```

```

Query Match          100.0%  Score 4029;  DB 5;  Length 770;
Best Local Similarity 100.0%  Pred. No. 0;
Matches 770: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 MAOMNOLQDLDTTRYKQLHQLYSDTFPMELROFLAPWIESODMAVAASKESHATLVFHNL 60
DB 1 MAOMNOLQDLDTTRYKQLHQLYSDTFPMELROFLAPWIESODMAVAASKESHATLVFHNL 60
OY 61 LGEIDDOYSRFLDOESNVLYOHNLRIKOFLOSRYLEKPMELIARIYAKCLIMESRLLOTAA 120
DB 61 LGEIDDOYSRFLDOESNVLYOHNLRIKOFLOSRYLEKPMELIARIYAKCLIMESRLLOTAA 120
OY 121 TAAOOGGQANHPTAAVVTTEKQOMLEOHLDQVRKRVODLEOKMKVVENLODDDFNFKTLK 180
DB 121 TAAOOGGQANHPTAAVVTTEKQOMLEOHLDQVRKRVODLEOKMKVVENLODDDFNFKTLK 180
OY 181 SOGDMODLNGNNSVTRQKMOOLEOMLTALDQMRISYSELAGLISAMEYVOKTLTDEEL 240
DB 181 SOGDMODLNGNNSVTRQKMOOLEOMLTALDQMRISYSELAGLISAMEYVOKTLTDEEL 240
OY 241 ADMKRRPEIACIGPPNICLDRLNNMTSLAESQLOTROQIKKLELOOKVSKGDPITVO 300
DB 241 ADMKRRPEIACIGPPNICLDRLNNMTSLAESQLOTROQIKKLELOOKVSKGDPITVO 300
OY 301 HRPMLEERIVELFRNLMSAFVVEROPCPMPHDPRLVIKTGVQFTTKYRLLVKPPELAY 360
DB 301 HRPMLEERIVELFRNLMSAFVVEROPCPMPHDPRLVIKTGVQFTTKYRLLVKPPELAY 360
OY 361 OLKIKYCIDKDSQVAALGSKRFNLGTNTKVMNEESNGLSAEFKHLTLREORCGN 420
DB 361 OLKIKYCIDKDSQVAALGSKRFNLGTNTKVMNEESNGLSAEFKHLTLREORCGN 420
OY 421 GGRANDASLIVTEELHLITFEFEVYHOGKIDLETHSLPYVVISITICOMPNAASILAY 480
DB 421 GGRANDASLIVTEELHLITFEFEVYHOGKIDLETHSLPYVVISITICOMPNAASILAY 480
OY 481 NMLTNNPKNVNFTKPPICITWDQVAEVLWSQFSSTTKRGLSIEQLTTLAEKLLPGPVNS 540
DB 481 NMLTNNPKNVNFTKPPICITWDQVAEVLWSQFSSTTKRGLSIEQLTTLAEKLLPGPVNS 540
OY 541 GCOITAAKCKEKEMAKGFSFWWMLNDIITDLYKYLALAMNEGYINGFISKEREALIST 600
DB 541 GCOITAAKCKEKEMAKGFSFWWMLNDIITDLYKYLALAMNEGYINGFISKEREALIST 600
OY 601 KPGTFLLPFSESSKEGGYTFTVVEKDISGKTQIOSVEPYTKOOLNNMSFAELIMGYKIM 660
DB 601 KPGTFLLPFSESSKEGGYTFTVVEKDISGKTQIOSVEPYTKOOLNNMSFAELIMGYKIM 660
OY 661 DATNIIIVSPLVLYPIPEKEAFAFKYCRPSOEHPREADPGSAAPYIKTFICVPTTCSN 720
DB 661 DATNIIIVSPLVLYPIPEKEAFAFKYCRPSOEHPREADPGSAAPYIKTFICVPTTCSN 720
OY 721 TIDLPMSPRTIOSLMQFGNNGEAPBSAGGOFESLTFPMDLTSECATSMP 770
DB 721 TIDLPMSPRTIOSLMQFGNNGEAPBSAGGOFESLTFPMDLTSECATSMP 770

```

```

DB 721 TIDLPMSPRTIOSLMQFGNNGEAPBSAGGOFESLTFPMDLTSECATSMP 770

```

```

RESULT 9
US-09-012-710-8
: Sequence 8, Application US/09012710
: Patent No. 6087478
: GENERAL INFORMATION:
: APPLICANT: Vinkemeier, Uwe
: APPLICANT: Moarefi, Ismail
: APPLICANT: Darnell, Jr., James E.
: APPLICANT: Kuriyan, John
: TITLE OF INVENTION: A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
: NUMBER OF SEQUENCES: 13
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Klauber & Jackson
: STREET: 411 Hackensack Avenue, 4th floor
: CITY: Hackensack
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07601
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/012,710
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Jackson Esq., David A.
: REGISTRATION NUMBER: 26,742
: REFERENCE/DOCKET NUMBER: 600-1-194
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201-487-5800
: TELEFAX: 201-343-1684
: TELEX: 133521
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 770 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHEICAL: NO
: US-09-012-710-8

```

```

Query Match          99.7%  Score 4018;  DB 3;  Length 770;
Best Local Similarity 99.7%  Pred. No. 0;
Matches 768: Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 MAOMNOLQDLDTTRYKQLHQLYSDTFPMELROFLAPWIESODMAVAASKESHATLVFHNL 60
DB 1 MAOMNOLQDLDTTRYKQLHQLYSDTFPMELROFLAPWIESODMAVAASKESHATLVFHNL 60
OY 61 LGEIDDOYSRFLDOESNVLYOHNLRIKOFLOSRYLEKPMELIARIYAKCLIMESRLLOTAA 120
DB 61 LGEIDDOYSRFLDOESNVLYOHNLRIKOFLOSRYLEKPMELIARIYAKCLIMESRLLOTAA 120
OY 121 TAAOOGGQANHPTAAVVTTEKQOMLEOHLDQVRKRVODLEOKMKVVENLODDDFNFKTLK 180
DB 121 TAAOOGGQANHPTAAVVTTEKQOMLEOHLDQVRKRVODLEOKMKVVENLODDDFNFKTLK 180
OY 181 SOGDMODLNGNNSVTRQKMOOLEOMLTALDQMRISYSELAGLISAMEYVOKTLTDEEL 240
DB 181 SOGDMODLNGNNSVTRQKMOOLEOMLTALDQMRISYSELAGLISAMEYVOKTLTDEEL 240
OY 241 ADMKRRPEIACIGPPNICLDRLNNMTSLAESQLOTROQIKKLELOOKVSKGDPITVO 300
DB 241 ADMKRRPEIACIGPPNICLDRLNNMTSLAESQLOTROQIKKLELOOKVSKGDPITVO 300

```

OY 301 HRPMLEERIVELFRNLMSAFVEROPCMHPRDLVIKTGVOTTKVRLLVKPEPLNY 360
DB 301 HRPMLEERIVELFRNLMSAFVEROPCMHPRDLVIKTGVOTTKVRLLVKPEPLNY 360
OY 361 OUKIKVCIDKDSGDVAALGSRKFNILGTNTKVMNNEESNNSLSAEFKHLTLRQRCGN 420
DB 361 OUKIKVCIDKDSGDVAALGSRKFNILGTNTKVMNNEESNNSLSAEFKHLTLRQRCGN 420
OY 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPVVVISNICOMPANAMASILMY 480
DB 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPVVVISNICOMPANAMASILMY 480
OY 481 NMLTNNPKVNFETKPRIGTMDQVAEVLISWOPSSITTKRGLSIEOLTTLAEKLLGPGVNS 540
DB 481 NMLTNNPKVNFETKPRIGTMDQVAEVLISWOPSSITTKRGLSIEOLTTLAEKLLGPGVNS 540
OY 5 1 GCOITWAKFCCKENMAGKGSFWVWLDNIIDLKYYILALMNGYIMGFISKERERAILST 600
DB 541 GCOITWAKFCCKENMAGKGSFWVWLDNIIDLKYYILALMNGYIMGFISKERERAILST 600
OY 601 KPPGTFLLRFSSSKREGVYFTWVEKDISGKTQIOSVEPYTKOOLNNSFAELIMGYKIM 660
DB 601 KPPGTFLLRFSSSKREGVYFTWVEKDISGKTQIOSVEPYTKOOLNNSFAELIMGYKIM 660
OY 661 DATNLIYSPVLYLPDIPKEFAFGKYCRPESQEHPRADGSAAPYLKTKFCVTPPTCSN 720
DB 661 DATNLIYSPVLYLPDIPKEFAFGKYCRPESQEHPRADGSAAPYLKTKFCVTPPTCSN 720
OY 721 TIDLPMSPRTLDLSLMQFNGNGGAEPSSAGOFESLTFDMDLTSECATSPM 770
DB 721 TIDLPMSPRTLDLSLMQFNGNGGAEPSSAGOFESLTFDMDLTSECATSPM 770

RESULT 10
US-09-364-970-5
: Sequence 5, Application US/09364970
: Patent No. 6235873
: GENERAL INFORMATION:
: APPLICANT: Bromberg, Jacqueline
: TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR
: TITLE OF INVENTION: USES FOR IDENTIFYING MODULATORS OF ACTIVITY INCLUDING
: FILE REFERENCE: 600-1-252
: CURRENT APPLICATION NUMBER: US/09/364,970
: CURRENT FILING DATE: 1999-07-31
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 5
: LENGTH: 770
: TYPE: PRT
: ORGANISM: Mus musculus
US-09-364-970-5

Query Match 99.7%: Score 4016; DB 4; Length 770;
Best Local Similarity 99.7%: Pred. No. 0;
Matches 768; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MAOMNLOLOLDRYKLOLQOLYSDPRMELROFLAPWISODMAVAASKESHATLVFHNL 60
DB 1 MAOMNLOLOLDRYKLOLQOLYSDPRMELROFLAPWISODMAVAASKESHATLVFHNL 60
OY 61 LGEIDDOYSRFIOESNVLYOHNLRIKOFLOSRYLEKPMELIARIVARCLMEESRLQTA 120
DB 61 LGEIDDOYSRFIOESNVLYOHNLRIKOFLOSRYLEKPMELIARIVARCLMEESRLQTA 120
OY 121 TAAOOGGQANHPRTAAVYTKOQMLEOHLODVRKRVODLEOKMKVVENLDDPFNFKTLK 180
DB 121 TAAOOGGQANHPRTAAVYTKOQMLEOHLODVRKRVODLEOKMKVVENLDDPFNFKTLK 180
OY 181 SOGDMDLNGNNSVTRKMOOLEOMLTALDQMRSTIVSELALLSAMEYVOGTLTDEEL 240
DB 181 SOGDMDLNGNNSVTRKMOOLEOMLTALDQMRSTIVSELALLSAMEYVOGTLTDEEL 240

DB 181 SOGDMDLNGNNSVTRKMOOLEOMLTALDQMRSTIVSELALLSAMEYVOGTLTDEEL 240
OY 241 ADMKRREIACIGGPNICIDLENNMITSIAESQOTROOIKKELEOQSYKGPDIYO 300
DB 241 ADMKRREIACIGGPNICIDLENNMITSIAESQOTROOIKKELEOQSYKGPDIYO 300
OY 301 HRPMLEERIVELFRNLMSAFVEROPCMHPRDLVIKTGVOTTKVRLLVKPEPLNY 360
DB 301 HRPMLEERIVELFRNLMSAFVEROPCMHPRDLVIKTGVOTTKVRLLVKPEPLNY 360
OY 361 OUKIKVCIDKDSGDVAALGSRKFNILGTNTKVMNNEESNNSLSAEFKHLTLRQRCGN 420
DB 361 OUKIKVCIDKDSGDVAALGSRKFNILGTNTKVMNNEESNNSLSAEFKHLTLRQRCGN 420
OY 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPVVVISNICOMPANAMASILMY 480
DB 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPVVVISNICOMPANAMASILMY 480
OY 481 NMLTNNPKVNFETKPRIGTMDQVAEVLISWOPSSITTKRGLSIEOLTTLAEKLLGPGVNS 540
DB 481 NMLTNNPKVNFETKPRIGTMDQVAEVLISWOPSSITTKRGLSIEOLTTLAEKLLGPGVNS 540
OY 541 GCOITWAKFCCKENMAGKGSFWVWLDNIIDLKYYILALMNGYIMGFISKERERAILST 600
DB 541 GCOITWAKFCCKENMAGKGSFWVWLDNIIDLKYYILALMNGYIMGFISKERERAILST 600
OY 601 KPPGTFLLRFSSSKREGVYFTWVEKDISGKTQIOSVEPYTKOOLNNSFAELIMGYKIM 660
DB 601 KPPGTFLLRFSSSKREGVYFTWVEKDISGKTQIOSVEPYTKOOLNNSFAELIMGYKIM 660
OY 661 DATNLIYSPVLYLPDIPKEFAFGKYCRPESQEHPRADGSAAPYLKTKFCVTPPTCSN 720
DB 661 DATNLIYSPVLYLPDIPKEFAFGKYCRPESQEHPRADGSAAPYLKTKFCVTPPTCSN 720
OY 721 TIDLPMSPRTLDLSLMQFNGNGGAEPSSAGOFESLTFDMDLTSECATSPM 770
DB 721 TIDLPMSPRTLDLSLMQFNGNGGAEPSSAGOFESLTFDMDLTSECATSPM 770

RESULT 11
US-08-416-581B-9
: Sequence 9, Application US/08416581B
: Patent No. 5719042
: GENERAL INFORMATION:
: APPLICANT: Kishimoto, Tadamiatsu
: APPLICANT: Akira, Shizuo
: TITLE OF INVENTION: TRANSCRIPTION FACTOR APRF
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
: STREET: 2100 Pennsylvania Avenue, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/416,581B
: FILING DATE: 04-APR-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 6-65825/1994
: FILING DATE: 04-APR-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Nakamura, Dean H.
: REGISTRATION NUMBER: 33,981
: REFERENCE/DOCKET NUMBER: O-37891
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)293-7060

TELEFAX: (202)293-7860
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 770 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-416-581B-9

Query Match 99.4%; Score 4006; DB 1; Length 770;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 765; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAQNNQLOQDLTRYLKQHLQSDTFPMELROFLAPWIESODMAVAASKESHATLVFHNL 60
DB 1 MAQNNQLOQDLTRYLKQHLQSDTFPMELROFLAPWIESODMAVAASKESHATLVFHNL 60
QY 61 LGEIDQOYSRFLQESNVLYOHNLRLIKOFLQSRYLEKPMELIARIYARCLMESSRLQTA 120
DB 61 LGEIDQOYSRFLQESNVLYOHNLRLIKOFLQSRYLEKPMELIARIYARCLMESSRLQTA 120
QY 121 TAAQGGGANHPPTAAVTEKQOMLEQHLQDVRRKVODLEQKKVVENLQDDDFENYKTLK 180
DB 121 TAAQGGGANHPPTAAVTEKQOMLEQHLQDVRRKVODLEQKKVVENLQDDDFENYKTLK 180
QY 181 SOGDMDLNGNNOYSTROKMOOLEQMLTALDOMRSTIVSELGLISAMEYVOKLTIDEEL 240
DB 181 SOGDMDLNGNNOYSTROKMOOLEQMLTALDOMRSTIVSELGLISAMEYVOKLTIDEEL 240
QY 241 ADMKRRPEIACIGPPNICLDRLENNITSLAESQLOTROQIKKLEELQOKVSKGDPVQ 300
DB 241 ADMKRRQOACIGPPNICLDRLENNITSLAESQLOTROQIKKLEELQOKVSKGDPVQ 300
QY 301 HRPMLERIVELFRLNLSAFYVEROPCMHPMDRPLVKTGVQFTTKRLLVKKPELNY 360
DB 301 HRPMLERIVELFRLNLSAFYVEROPCMHPMDRPLVKTGVQFTTKRLLVKKPELNY 360
QY 361 QLKIKVCIDKDSGDVAALGSRKFNLGNTKVMNNEESNNGSLSAEFKHLTLREORGN 420
DB 361 QLKIKVCIDKDSGDVAALGSRKFNLGNTKVMNNEESNNGSLSAEFKHLTLREORGN 420
QY 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPVVVISNICOMPNAASILMY 480
DB 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPVVVISNICOMPNAASILMY 480
QY 481 NMLTNPKVNFETPRPIGTMDOVAEVLISMOFSSPTTKRGSLTEOLTTLAEXLLGCVNYS 540
DB 481 NMLTNPKVNFETPRPIGTMDOVAEVLISMOFSSPTTKRGSLTEOLTTLAEXLLGCVNYS 540
QY 541 GCOITWAFCKENMAKGFSFWMLDNIIDLVKYLLALAMNGYIMGFTSKRERAILST 600
DB 541 GCOITWAFCKENMAKGFSFWMLDNIIDLVKYLLALAMNGYIMGFTSKRERAILST 600
QY 601 KPPGFFLLRFSSSSKEGCVTFTWVEKDISGKTQIOSVEYTTKOOLNMFPAELIMCYKIM 660
DB 601 KPPGFFLLRFSSSSKEGCVTFTWVEKDISGKTQIOSVEYTTKOOLNMFPAELIMCYKIM 660
QY 661 DATNIIIVSLVLYPDIPKEEAFGYKRCRESOEHPEADGSAAPYLKTFICVPTTCSN 720
DB 661 DATNIIIVSLVLYPDIPKEEAFGYKRCRESOEHPEADGSAAPYLKTFICVPTTCSN 720
QY 721 TIDLPMSPRTLDLSLMQFGNNGGAEPASAGQESLTFMDLTSECATSPM 770
DB 721 TIDLPMSPRTLDLSLMQFGNNGGAEPASAGQESLTFMDLTSECATSPM 770

RESULT 12
US-08-416-581B-1
: Sequence 1, Application US/08416581B
: Patent No. 5719042
: GENERAL INFORMATION:
: APPLICANT: Kishimoto, Tadamitsu

APPLICANT: Akira, Shizuo
: TITLE OF INVENTION: TRANSCRIPTION FACTOR APRF
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SUGHRUE, MIOM, ZINN, MACPEAK & SEAS
: STREET: 2100 Pennsylvania Avenue, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/416,581B
: FILING DATE: 04-APR-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 6-65825/1994
: FILING DATE: 04-APR-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Nakamura, Dean H.
: REGISTRATION NUMBER: 33,981
: REFERENCE/DOCKET NUMBER: 0-37891
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)293-7060
: TELEFAX: (202)293-7860
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 770 amino acids
: TYPE: amino acid
: STRANDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-416-581B-1

Query Match 98.7%; Score 3977; DB 1; Length 770;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 759; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 MAQNNQLOQDLTRYLKQHLQSDTFPMELROFLAPWIESODMAVAASKESHATLVFHNL 60
DB 1 MAQNNQLOQDLTRYLKQHLQSDTFPMELROFLAPWIESODMAVAASKESHATLVFHNL 60
QY 61 LGEIDQOYSRFLQESNVLYOHNLRLIKOFLQSRYLEKPMELIARIYARCLMESSRLQTA 120
DB 61 LGEIDQOYSRFLQESNVLYOHNLRLIKOFLQSRYLEKPMELIARIYARCLMESSRLQTA 120
QY 121 TAAQGGGANHPPTAAVTEKQOMLEQHLQDVRRKVODLEQKKVVENLQDDDFENYKTLK 180
DB 121 TAAQGGGANHPPTAAVTEKQOMLEQHLQDVRRKVODLEQKKVVENLQDDDFENYKTLK 180
QY 181 SOGDMDLNGNNOYSTROKMOOLEQMLTALDOMRSTIVSELGLISAMEYVOKLTIDEEL 240
DB 181 SOGDMDLNGNNOYSTROKMOOLEQMLTALDOMRSTIVSELGLISAMEYVOKLTIDEEL 240
QY 241 ADMKRRQOACIGPPNICLDRLENNITSLAESQLOTROQIKKLEELQOKVSKGDPVQ 300
DB 241 ADMKRRQOACIGPPNICLDRLENNITSLAESQLOTROQIKKLEELQOKVSKGDPVQ 300
QY 301 HRPMLERIVELFRLNLSAFYVEROPCMHPMDRPLVKTGVQFTTKRLLVKKPELNY 360
DB 301 HRPMLERIVELFRLNLSAFYVEROPCMHPMDRPLVKTGVQFTTKRLLVKKPELNY 360
QY 361 QLKIKVCIDKDSGDVAALGSRKFNLGNTKVMNNEESNNGSLSAEFKHLTLREORGN 420
DB 361 QLKIKVCIDKDSGDVAALGSRKFNLGNTKVMNNEESNNGSLSAEFKHLTLREORGN 420
QY 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPVVVISNICOMPNAASILMY 480
DB 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPVVVISNICOMPNAASILMY 480

Db 421 GGRANDASLIYTEELHLTFETEVYHOGKIDLETHSLSVVVISNICMPNAMSILMY 480
OY 481 NMLTNNPKVNFPTKPPICITWQOVAEVLWSQFSTTKRGSLIDOLTTTAEKLLGPGVNS 540
Db 481 NMLTNNPKVNFPTKPPICITWQOVAEVLWSQFSTTKRGSLIDOLTTTAEKLLGPGVNS 540
OY 541 GCOITWAKCKENMAGKGFSPVWMLDNIIDLKRYTLALMNEGYIMGFISKEHERAIIIST 600
Db 541 GCOITWAKCKENMAGKGFSPVWMLDNIIDLKRYTLALMNEGYIMGFISKEHERAIIIST 600
OY 601 KPPGFTLLRFSESSKEGVTFTWVKEDISGKTQIOSVEPYTKOOLNNMSFAEIIIMGYKIM 660
Db 601 KPPGFTLLRFSESSKEGVTFTWVKEDISGKTQIOSVEPYTKOOLNNMSFAEIIIMGYKIM 660
OY 661 DATNIIIVSLVLYLPDIPKEEAFGKYCRPSOEHPRADGSAAPYKTFICVTPPTCSN 720
Db 661 DATNIIIVSLVLYLPDIPKEEAFGKYCRPSOEHPRADGSAAPYKTFICVTPPTCSN 720
OY 721 TIDLPMSPRTLDLSLMQFGNNGGAEPSAGGFESLTFDMDLTSECATSPM 770
Db 721 TIDLPMSPRTLDLSLMQFGNNGGAEPSAGGFESLTFDMDLTSECATSPM 770

RESULT 13
US-08-416-581B-5
: Sequence 5, Application US/08416581B
: Patent No. 5719042
: GENERAL INFORMATION:
: APPLICANT: Kishimoto, Tadamitsu
: APPLICANT: Akita, Shizuo
: TITLE OF INVENTION: TRANSCRIPTION FACTOR APRE
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
: STREET: 2100 Pennsylvania Avenue, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/416,581B
: FILING DATE: 04-APR-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 6-65825/1994
: FILING DATE: 04-APR-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Nakamura, Dean H.
: REGISTRATION NUMBER: 33,981
: REFERENCE/DOCKET NUMBER: O-37891
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)293-7060
: TELEFAX: (202)293-7860
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 770 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-416-581B-5

Query Match 98.7%; Score 3977; DB 1; Length 770;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 759; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
OY 1 MAOMNLOQLODTRVYLKQLHOLVSDTFPMELRQFLAPWIESQDAWYAASKESHATLVFHNL 60
|||||

Db 1 MAOMNLOQLODTRVYLKQLHOLVSDTFPMELRQFLAPWIESQDAWYAASKESHATLVFHNL 60
OY 61 LCEIDOOVSRLFOESNVLYOHNLRLRIKQFLQSRYLEKPMELIARIYARCLMESSRLLOTA 120
Db 61 LCEIDOOVSRLFOESNVLYOHNLRLRIKQFLQSRYLEKPMELIARIYARCLMESSRLLOTA 120
OY 121 TAAOOGGQANHPTAAVVFTEKQOMLEOHLODVKRPRODI.EOKKVVENLODDDFPNYKTLK 180
Db 121 TAAOOGGQANHPTAAVVFTEKQOMLEOHLODVKRPRODLEOKKVVENLODDDFPNYKTLK 180
OY 181 SOGDMODLGNNOVYTRQKMOOLEQMLTALDQMRISIVSELAGLISAMEYVOKTLTDEEL 240
Db 181 SOGDMODLGNNOVYTRQKMOOLEQMLTALDQMRISIVSELAGLISAMEYVOKTLTDEEL 240
OY 241 ADMKRRPEIACIGSPNLCIDLENNITSLASOLQOTROQIKKLELOOKVSKDPIYQ 300
Db 241 ADMKRRPEIACIGSPNLCIDLENNITSLASOLQOTROQIKKLELOOKVSKDPIYQ 300
OY 301 HRPMLEERIVELFRNLMSAFYVEROPCMHPDRPI.VIKTGVQFTTKVRL.VKPEELNY 360
Db 301 HRPMLEERIVELFRNLMSAFYVEROPCMHPDRPI.VIKTGVQFTTKVRL.VKPEELNY 360
OY 361 QLKIKYCIDKSGDVAALGSKRFNLGTNTKYMMNEESNNSLSAEFKILTLREORGN 420
Db 361 QLKIKYCIDKSGDVAALGSKRFNLGTNTKYMMNEESNNSLSAEFKILTLREORGN 420
OY 421 GGRANDASLIYTEELHLTFETEVYHOGKIDLETHSLPVVVISNICMPNAMSILMY 480
Db 421 GGRANDASLIYTEELHLTFETEVYHOGKIDLETHSLSVVVISNICMPNAMSILMY 480
OY 481 NMLTNNPKVNFPTKPPICITWQOVAEVLWSQFSTTKRGSLIDOLTTTAEKLLGPGVNS 540
Db 481 NMLTNNPKVNFPTKPPICITWQOVAEVLWSQFSTTKRGSLIDOLTTTAEKLLGPGVNS 540
OY 541 GCOITWAKCKENMAGKGFSPVWMLDNIIDLKRYTLALMNEGYIMGFISKEHERAIIIST 600
Db 541 GCOITWAKCKENMAGKGFSPVWMLDNIIDLKRYTLALMNEGYIMGFISKEHERAIIIST 600
OY 601 KPPGFTLLRFSESSKEGVTFTWVKEDISGKTQIOSVEPYTKOOLNNMSFAEIIIMGYKIM 660
Db 601 KPPGFTLLRFSESSKEGVTFTWVKEDISGKTQIOSVEPYTKOOLNNMSFAEIIIMGYKIM 660
OY 661 DATNIIIVSLVLYLPDIPKEEAFGKYCRPSOEHPRADGSAAPYKTFICVTPPTCSN 720
Db 661 DATNIIIVSLVLYLPDIPKEEAFGKYCRPSOEHPRADGSAAPYKTFICVTPPTCSN 720
OY 721 TIDLPMSPRTLDLSLMQFGNNGGAEPSAGGFESLTFDMDLTSECATSPM 770
Db 721 TIDLPMSPRTLDLSLMQFGNNGGAEPSAGGFESLTFDMDLTSECATSPM 770

RESULT 14
US-09-087-465-6
: Sequence 6, Application US/09087465A
: Patent No. 6160992
: GENERAL INFORMATION:
: APPLICANT: Vinkemeier, Uwe
: APPLICANT: Chen, Xiaomin
: APPLICANT: Darnell Jr., James E
: TITLE OF INVENTION: A CRYSTAL OF THE CORE PORTION OF A STAT AND METHODS OF
: FILE REFERENCE: 600-1-229
: CURRENT APPLICATION NUMBER: US/09/087,465A
: FILING DATE: 1998-05-29
: NUMBER OF SEQ ID NOS: 37
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 6
: LENGTH: 770
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-087-465-6

Query Match 98.7%; Score 3977; DB 4; Length 770;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 759; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

```

OY 1 MAOMNLOOQDTRTYLQOLHOLYSDTFPMELROFLAPWIESQDMAYAKSKESHATLVFHNH 60
    |||
DB 1 MAOMNLOOQDTRTYLQOLHOLYSDTFPMELROFLAPWIESQDMAYAKSKESHATLVFHNH 60
    |||
OY 61 LGEDIDQYSFLEDSNVLYOHNLRIKOPLOSRYLEKPMELIARIYARCLMEESRLLOTAA 120
    |||
DB 61 LGEDIDQYSFLEDSNVLYOHNLRIKOPLOSRYLEKPMELIARIYARCLMEESRLLOTAA 120
    |||
OY 121 TAAOOGGOANHPPTAAVYTEKQOMLEOHLQDVRKRVODLEOKMKVENVLQDDPFNYKTKL 180
    |||
DB 121 TAAOOGGOANHPPTAAVYTEKQOMLEOHLQDVRKRVODLEOKMKVENVLQDDPFNYKTKL 180
    |||
OY 181 SOGDMDLNGNNSVTRKQMOOLEOMLTALDOKRSTVSELAGLSAMEYVOKTLTDEEL 240
    |||
DB 181 SOGDMDLNGNNSVTRKQMOOLEOMLTALDOKRSTVSELAGLSAMEYVOKTLTDEEL 240
    |||
OY 241 ADMKRRPEIACIGPPNICDLRLNNITSLAESOLQTRQOIKKLEELQOKVSYKGDPIYO 300
    |||
DB 241 ADMKRRPEIACIGPPNICDLRLNNITSLAESOLQTRQOIKKLEELQOKVSYKGDPIYO 300
    |||
OY 301 HRPMLEERIVELFRNLKMSAFVVEROPCMHPDRPLVITGYOFTTKVALLKFPPELNY 360
    |||
DB 301 HRPMLEERIVELFRNLKMSAFVVEROPCMHPDRPLVITGYOFTTKVALLKFPPELNY 360
    |||
OY 361 QLKIKVCIDKSGDVAALNGSRKFNLGTNTKYMMNEESNNGSLSAEFKHLTREORCGN 420
    |||
DB 361 QLKIKVCIDKSGDVAALNGSRKFNLGTNTKYMMNEESNNGSLSAEFKHLTREORCGN 420
    |||
OY 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPLVYVINSICOMPANASILMY 480
    |||
DB 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPLVYVINSICOMPANASILMY 480
    |||
OY 481 NMLTNNPKVNFPTKPRIGTMDOVAEVLWSQFSSTTKRGISTEOLTTLAELKLGPGVNY 540
    |||
DB 481 NMLTNNPKVNFPTKPRIGTMDOVAEVLWSQFSSTTKRGISTEOLTTLAELKLGPGVNY 540
    |||
OY 541 GCOITMANCKENMAGKGSFWWMLDNIIDLVKYILALMNEGYIMGFISKEBERA1LST 600
    |||
DB 541 GCOITMANCKENMAGKGSFWWMLDNIIDLVKYILALMNEGYIMGFISKEBERA1LST 600
    |||
OY 601 KPPGTFLLRFSSSKEGVTFTEWEDISGKTQIOSVEPYTKOOLNMSFAELIMGYKIM 660
    |||
DB 601 KPPGTFLLRFSSSKEGVTFTEWEDISGKTQIOSVEPYTKOOLNMSFAELIMGYKIM 660
    |||
OY 661 DATNILVSPLYLYLPDIPEEAFGKYCRPESOEHPREADGSAPYLKTFICVTPPTCSN 720
    |||
DB 661 DATNILVSPLYLYLPDIPEEAFGKYCRPESOEHPREADGSAPYLKTFICVTPPTCSN 720
    |||
OY 721 TIDLPMSPRALDLSLMQFGNNGEAGPSAGQFESLTFDMDLTSECATSPM 770
    |||
DB 721 TIDLPMSPRALDLSLMQFGNNGEAGPSAGQFESLTFDMDLTSECATSPM 770
    |||

```

```

; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,099A
; FILING DATE: 15-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard Aron
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-59451-1/BAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
US-08-276-099A-14

```

Query Match 98.7%; Score 3977; DB 1; Length 771;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 759; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

```

OY 1 MAOMNLOOQDTRTYLQOLHOLYSDTFPMELROFLAPWIESQDMAYAKSKESHATLVFHNH 60
    |||
DB 1 MAOMNLOOQDTRTYLQOLHOLYSDTFPMELROFLAPWIESQDMAYAKSKESHATLVFHNH 60
    |||
OY 61 LGEDIDQYSFLEDSNVLYOHNLRIKOPLOSRYLEKPMELIARIYARCLMEESRLLOTAA 120
    |||
DB 61 LGEDIDQYSFLEDSNVLYOHNLRIKOPLOSRYLEKPMELIARIYARCLMEESRLLOTAA 120
    |||
OY 121 TAAOOGGOANHPPTAAVYTEKQOMLEOHLQDVRKRVODLEOKMKVENVLQDDPFNYKTKL 180
    |||
DB 121 TAAOOGGOANHPPTAAVYTEKQOMLEOHLQDVRKRVODLEOKMKVENVLQDDPFNYKTKL 180
    |||
OY 181 SOGDMDLNGNNSVTRKQMOOLEOMLTALDOKRSTVSELAGLSAMEYVOKTLTDEEL 240
    |||
DB 181 SOGDMDLNGNNSVTRKQMOOLEOMLTALDOKRSTVSELAGLSAMEYVOKTLTDEEL 240
    |||
OY 241 ADMKRRPEIACIGPPNICDLRLNNITSLAESOLQTRQOIKKLEELQOKVSYKGDPIYO 300
    |||
DB 241 ADMKRRPEIACIGPPNICDLRLNNITSLAESOLQTRQOIKKLEELQOKVSYKGDPIYO 300
    |||
OY 301 HRPMLEERIVELFRNLKMSAFVVEROPCMHPDRPLVITGYOFTTKVALLKFPPELNY 360
    |||
DB 301 HRPMLEERIVELFRNLKMSAFVVEROPCMHPDRPLVITGYOFTTKVALLKFPPELNY 360
    |||
OY 361 QLKIKVCIDKSGDVAALNGSRKFNLGTNTKYMMNEESNNGSLSAEFKHLTREORCGN 420
    |||
DB 361 QLKIKVCIDKSGDVAALNGSRKFNLGTNTKYMMNEESNNGSLSAEFKHLTREORCGN 420
    |||
OY 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPLVYVINSICOMPANASILMY 480
    |||
DB 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPLVYVINSICOMPANASILMY 480
    |||
OY 481 NMLTNNPKVNFPTKPRIGTMDOVAEVLWSQFSSTTKRGISTEOLTTLAELKLGPGVNY 540
    |||
DB 481 NMLTNNPKVNFPTKPRIGTMDOVAEVLWSQFSSTTKRGISTEOLTTLAELKLGPGVNY 540
    |||
OY 541 GCOITMANCKENMAGKGSFWWMLDNIIDLVKYILALMNEGYIMGFISKEBERA1LST 600
    |||
DB 541 GCOITMANCKENMAGKGSFWWMLDNIIDLVKYILALMNEGYIMGFISKEBERA1LST 600
    |||

```

Tue Mar 19 16:20:11 2002

us-08-212-185-12.raii

Page 12

Oy	601	KPPGFFLLRFSSSSNEGCVFTTWEXKDISCKTOIQSEVPTKQOOLNMSAEIIMGKIM	660
Db	601	KPPGFFLLRFSSSSNEGCVFTTWEXKDISCKTOIQSEVPTKQOOLNMSAEIIMGKIM	660
Oy	661	DATILNLSPVLYIKVDPDIKEAFKCKYRPSOEHEHPADPDSAAPLYKTRICVPTTCSN	720
Db	661	DATILNLSPVLYIKVDPDIKEAFKCKYRPSOEHEHPADPDSAAPLYKTRICVPTTCSN	720
Oy	661	DATILNLSPVLYIKVDPDIKEAFKCKYRPSOEHEHPADPDSAAPLYKTRICVPTTCSN	720
Db	661	DATILNLSPVLYIKVDPDIKEAFKCKYRPSOEHEHPADPDSAAPLYKTRICVPTTCSN	720
Oy	721	TIDLPMSFRTIDSLMOFGNNGEGAEPSAGGOFESLTFDMDITJSECATSPM	770
Db	721	TIDLPMSFRTIDSLMOFGNNGEGAEPSAGGOFESLTFDMDITJSECATSPM	770

Search completed: March 19, 2002, 15:22:45
Job time: 3132 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 19, 2002, 14:01:18 ; Search time 38.35 Seconds
(without alignments)
1487.260 Million cell updates/sec

Title: US-08-212-185-12

Perfect score: 4029
Sequence: 1 MAQNNQLOQLDTRYLKQLHQ.....QFSLTFPMDLTSECATSPM 770

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4029	100.0	770	16	AA72082
2	4029	100.0	770	17	AAW03176
3	4018	99.7	770	21	AA12317
4	4006	99.4	770	16	AA822995
5	4006	99.4	770	20	AA03768
6	4006	99.4	793	21	AA58442
7	3977	98.7	770	22	AA19964
8	3970	98.5	770	16	AA82293
9	2087	51.8	393	22	AA72842
10	2017	50.1	750	16	AA72078
11	2017	50.1	750	17	AAW03168

12	2017	50.1	750	19	AAW62994	Human Stat1-alpha
13	2017	50.1	750	22	AA81962	Human signal trans
14	2017	50.1	786	21	AA854098	Human pancreatic C
15	1997	49.6	712	16	AA72079	Human Stat84. Hom
16	1997	49.6	712	17	AAW03170	Human STAT1-beta.
17	1997	49.6	712	19	AAW62995	Human Stat1-beta P
18	1984	49.2	749	16	AA72080	Mouse Stat1 (Stat9
19	1984	49.2	749	17	AAW03172	Mouse STAT1. Mus
20	1984	49.2	749	21	AA812376	N-terminal domain
21	1971.5	48.9	739	14	AA841334	91 KD ISGF-3alpha.
22	1945.5	48.3	701	14	AA841335	84 KD ISGF-3alpha.
23	1798	44.6	748	17	AAW01101	Signal transducer
24	1798	44.6	748	22	AA819965	Human signal trans
25	1786	44.3	704	17	AAW01102	Signal transducer
26	1780.5	44.2	749	21	AA812378	N-terminal domain
27	1716	42.6	748	16	AAW03174	Mouse STAT3. Mus
28	1714	42.5	748	16	AA72081	Mouse Stat4 (13st1
29	1648	40.9	582	16	AAW62996	Human truncated St
30	1388	34.5	271	22	AAV72841	Mouse Stat3 protei
31	1293	32.1	851	14	AA841333	113 KD ISGF-3alpha
32	1293	32.1	851	16	AA841337	Recognition factor
33	1293	32.1	851	17	AAW03166	Human STAT2. Homo
34	1293	32.1	851	21	AA812374	N-terminal domain
35	1293	32.1	851	22	AA819963	Human signal trans
36	1290	32.0	252	22	AAV72846	Mouse Stat3 protei
37	1212	30.1	236	22	AAV72847	Mouse Stat3 protei
38	1172	29.1	229	22	AAV72850	Mouse Stat3 protei
39	1168	29.0	229	22	AAV72853	Mouse Stat3 protei
40	1167	29.0	229	22	AAV72862	Mouse Stat3 protei
41	1157.5	28.7	228	22	AAV72861	Mouse Stat3 protei
42	1146	28.4	229	22	AAV72860	Mouse Stat3 protei
43	1142	28.3	223	22	AAV72854	Mouse Stat3 protei
44	1094	27.2	213	22	AAV72851	Mouse Stat3 protei
45	1049	26.0	376	22	AAV72845	Mouse Stat1 protei

ALIGNMENTS

RESULT 1	
AA72082	AA72082 standard; Protein: 770 AA.
XX	XX
AA72082:	
XX	XX
27-SEP-1995 (first entry)	
XX	XX
Mouse Stat3 (19sf6).	
XX	XX
Signal transducer and activator of transcription; STAT; 19sf6;	
KW	KW
Stat3: receptor recognition factor; transcription factor;	
KW	KW
cellular dephosphorylation; derangement; dysfunction;	
KW	KW
interferon-gamma.	
XX	XX
Mus sp.	
XX	XX
WO9508629-A.	
XX	XX
30-MAR-1995.	
XX	XX
26-SEP-1994;	94WO-US10849.
XX	XX
24-SEP-1993;	93US-0126588.
PR	PR
24-SEP-1993;	93US-0126595.
PR	PR
11-MAR-1994;	94US-0212184.
PR	PR
11-MAR-1994;	94US-0212185.
XX	XX
(UTRQ) UNIV ROCKEFELLER.	
PA	PA
XX	XX
Darnell JE, Schindler CW, Shuai K, Wen Z, Zhong Z;	
PI	PI
XX	XX
WPI: 1995 139598/18.	
DR	DR
N-PSDB: A 389340	

xx Receptor recognition factor implicated in transcriptional
 PT stimulation of genes - useful in drug screening assays and/or
 PT for treating cellular debilitations, derangements and/or
 PT dysfunction, etc.
 PS Claim 1; Page 107-110; 160pp; English.
 xx A fragment encoding the human Stat91 protein was used to screen a
 CC murine thymus and spleen cDNA for homologous proteins. A highly
 CC homologous gene (given in AA089338) was isolated that encoded a
 CC 91 kDa protein (AA072080) (Stat1) that was responsive to interferon-
 CC gamma. Using a fragment of the mouse gene as probe, 2 additional
 CC members of the 113-91 family of receptor recognition factor
 CC proteins were isolated. The 2 genes (AA089339-40) were cloned
 CC in plasmids 13sf1 and 19sf6 and encoded proteins termed Stat4
 CC (AA072081) and Stat3 (AA072082), respectively.
 xx
 SQ Sequence 770 AA;

Query Match 100.0%; Score 4029; DB 16; Length 770;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAOMNLOQLDRLRYLKHQLYSDFPPELRLQFLAPWIESQPMAYAAKESHATLVFHNL 60
 DB 1 maqmwqlqldirylkqhqlysdftfpmelrlqlapwiesqgwajaaakeshatlvfhn 60
 QY 61 LGEIDQVSRFLOESNVLYOHNLRRIKOFLQGRYLEKPEIARIYARCLWESRLLOTAA 120
 DB 61 lgeidqysrflgesnvlyqhnllrrikqflqsrylekpeiarivarclwessrlloata 120
 QY 121 TAAOOGGANPHTAVNTEKQOMLEOHLODYRKRVODLEQKKKVVENLODDDFNYKTK 180
 DB 121 taaogqganphtaavnteekqomleohlodvyrkrvodeqkkkvvendoddffnyktk 180
 QY 121 caaqqgqanphtaavnteekqomleohlodvyrkrvodeqkkkvvendoddffnyktk 180
 DB 121 caaqqgqanphtaavnteekqomleohlodvyrkrvodeqkkkvvendoddffnyktk 180
 QY 181 SOGDMDLNGNNSVTRKQMOLEOMLTALDOMRRISYSELGSLSAMEYOKTLDEEL 240
 DB 181 sqgdmdlgnngsvtrkqmqleqmltalqdmrrisvselgslsameyoktildedel 240
 QY 181 sqgdmdlgnngsvtrkqmqleqmltalqdmrrisvselgslsameyoktildedel 240
 DB 181 sqgdmdlgnngsvtrkqmqleqmltalqdmrrisvselgslsameyoktildedel 240
 QY 241 ADWKRREPIACIGPPNICLDRLENWITSLSAQSLOTROQIKKLELQKVSYGDPVQ 300
 DB 241 adwkrrepiactigppnicldrlenwitslsaesqlqtrqikklelqkvsyqgdpvq 300
 QY 301 HRPMLSEERIVELFRMLKSAFVERQPCMPHDPRLVKTGVOFTTKVRLKVPPELNY 360
 DB 301 hrpmlseerivelfrmlksafverqpcmpmhdpriylvktgvttkvrvllvppelny 360
 QY 361 QLKTKVCLDKSDGVAAALRGSRKFNILQTNKVMNNEESNNGSLAEFKHLTLRQRCGN 420
 DB 361 qlktkvcldksdgvaaalrgsrkfnllgtncvmneesngslaeekhlrlrqrcgn 420
 QY 421 GGRANCDASLVTEELHITFEYVHOGKLIDLETHSLPVVYVINSICOMPANASIMLY 480
 DB 421 ggrancdaslvteelhitefevhogklidlethslpvvyvinsicompnasimly 480
 QY 421 ggrancdaslvteelhitefevhogklidlethslpvvyvinsicompnasimly 480
 DB 421 ggrancdaslvteelhitefevhogklidlethslpvvyvinsicompnasimly 480
 QY 481 NMLTNPNKVNPFETPPICGTDQVAEVLNMOFSSTTKRGLSTEOLTTLAEXLLGGVNS 540
 DB 481 nmltnpnkvnpfetppicgtwdqvaevlnswfssctkrqlsteqltlaekllpgvny 540
 QY 541 GCQITWAKFCKENNAKGFSFWMLDNIIDLKXYILALMNGYIMGFISKEERAILST 600
 DB 541 gcqitwakfckennakgfsfwvldnidlykxyilalmngyimgfiskerailst 600
 QY 601 KPRCFILRSESSKEGCVTFMWKDLISGKTQIOSVEPYTQOQNNNSFAITMGYKIM 660
 DB 601 kprcfilrseesskegcvtfmwkdlisgktqiosvepytqoqnnnsfaimgykim 660
 QY 661 DATNILVSPLYLYPDJPKBEAFGKCRPESQHPADPGSAAPYLKTKFICVPTPTCSN 720
 DB 661 datnilvsplylypdjpkbeafgkcrpesqhpadpgsaapylktkficvptptcsn 720

QY 721 TIDPMSPTRLDLMOFNNGECAPSGCOFSELTFFDMLTSECATSPM 770
 DB 721 tidpmsptrlsdlmofnngecapsgcofseltffdmlltsecatspm 770

RESULT 2

AAW03176
 ID AAW03176 standard; Protein; 770 AA.

AC AAW03176;

DT 24-OCT-1996 (first entry)

DE Mouse STAT4.

KW STAT; STAT4; signal transducer and activator of transcription;

KW DNA binding protein; ligand; receptor; oncogenesis; inflammation;

OS Mus sp.

PH Key Location/Qualifiers

FT Domain 398..508

FT /label= DNA_binding_domain

PN W09620954-A2.

PD 11-JUL-1996.

PF 28-DEC-1995; 95WO-US17025.

PR 06-JAN-1995; 95US-0369796.

PA (UYRO) IV ROCKEFELLER.

PI Darnell JE, Horvath CM, Wen Z, Zhong Z;

DR WPI: 1996-333941/33.

DR N-PSDB; AAT31280.

PT New STAT protein DNA-binding domain peptide(s) - useful for

PT diagnosis, preventing or treating cellular dysfunction, e.g.

PS oncogenesis, inflammation, parasitic disease or autoimmunity

PS Disclosure: Page 87-90; 138pp; English.

CC Mouse signal transducer and activator of transcription (STAT)

CC protein STAT4 (AAW03176) serves a dual purpose, i.e. signal

CC transduction from ligand-activated receptor kinase complexes

CC followed by nuclear translocation and DNA binding to activate

CC transcription. Recombinant STAT4 can be obtd. using cDNA clone

CC 19sf6 (AAT31278) obtd. from splenic/thymic cells. STAT4 includes

CC a DNA-binding domain (see also AAW03167) capable of both

CC receptor recognition and message delivery via DNA binding in a

CC binding domains (see also AAW03165-75) are useful for screening

CC antagonists used to inhibit STAT-mediated signal transduction

CC and activation of transcription.

SQ Sequence 770 AA;

Query Match 100.0%; Score 4029; DB 17; Length 770;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAOMNLOQLDRLRYLKHQLYSDFPPELRLQFLAPWIESQPMAYAAKESHATLVFHNL 60
 DB 1 maqmwqlqldirylkqhqlysdftfpmelrlqlapwiesqgwajaaakeshatlvfhn 60
 QY 61 LGEIDQVSRFLOESNVLYOHNLRRIKOFLQGRYLEKPEIARIYARCLWESRLLOTAA 120
 DB 61 lgeidqysrflgesnvlyqhnllrrikqflqsrylekpeiarivarclwessrlloata 120


```

Db      361 qlikvcidkdsqdvaaalrgsrkfnlglntkvmmeesnglsjaefkhlilreqrcgn 420
OY      421 GGRANCDASLIYTEELHLITFETEVYHOGKIDLETHSLPVAVINSICOMPANASILMY 480
        |||||||
Db      421 ggrancdasliyteelhlitfetevyhgkikidlethslpvvniicompawasilwy 480
OY      481 NMLTNPKNVNFFTKPRPGTMDVAEVLISWOFSSSTTKRGLSIEQLTTLAEKLLGCVNYS 540
        |||||||
Db      481 nmltnpkvnfftkprpgtmdvaevliswofsssttkrglsieqlttlaekllpgvnyys 540
OY      541 GCOITWAKFCENMAGKGSFWVWLDNIIDLKYYILALMNGGYIMGFISKREERAILST 600
        |||||||
Db      541 gcoitwakfckenmagkgsfwwldnidlvkkyilalwmegylimgfiskerailstc 600
OY      601 KPPGFTLLRFSSSKEGVFTWVERKDISGKTQIQSVEPYTRKOOLNNMSFAEIMGYKIM 660
        |||||||
Db      601 kppgftllrfssskegyftfwvekdlsqkctqigsvepytkqqlnmsfaeilmgylim 660
OY      661 DATNIIIVSPVLYLPDIPKEEAFGKYCRPESQHPHADGSAAPYIKTKFICVPTTCSN 720
        |||||||
Db      661 datniiivsplylvpdipkeeafigkycrpesqhpdaadpysaapyiktkficvptlcsn 720
OY      721 TIDLPMSPRTLDLSLMQFCNNGCAEPSAGQFESLTFDMDLTSECATSPM 770
        |||||||
Db      721 tidlpmsprtldslmqfngngeaepsagqfesltfdmdltsecatspm 770

```

```

RESULT 4
AAR82995
ID      AAR82995 standard: Protein: 770 AA.
AC      AAR82995:
XX      25-MAR-1996 (first entry)
XX      Mouse liver acute phase response factor.
DE      Mouse liver acute phase response factor.
XX      Mouse: acute phase response factor: transcription factor;
KW      Interleukin-6; signal transmission; liver: antibody; antisense;
KW      ribozyme; antiinflammatory; antitumor; hypotensive; therapy.
OS      Mus musculus.
XX      EP676469-A2.
XX      11-OCT-1995.
PD      29-MAR-1995: 95EP-0104670.
XX      04-APR-1994: 94JP-0065825.
XX      (KISH/) KISHIMOTO T.
XX      PI Akira S, Kishimoto T;
XX      WPI: 1995-346089/45.
DR      N-PSDB: AAT05619.
XX      New acute phase response factor - for developing inhibitory agents
PT      for treating diseases induced by cytokine(s) such as IL-6, e.g.
XX      inflammatory diseases
XX      Claim 10; Page 20-22; 31pp: English.
XX      The sequence represents a mouse acute phase response factor (APRF),
XX      a transcription factor related to signal transmission of
XX      interleukin-6 (IL-6). The protein is encoded by a cDNA, isolated
XX      from a mouse liver cDNA library using a polymerase chain reaction
XX      product (amplified using primers derived from an IL-6-treated mouse
XX      liver peptide) as a probe. APRF-inhibitors, e.g. antibodies,
XX      antisense oligonucleotides or ribozymes, may be used to treat
XX      diseases induced by IL-6, e.g. inflammatory disease, leukemia,
XX      cancer, osteoclastia, pulmonary hypertension, etc.

```

```

XX      SQ      Sequence      770 AA:
OY      1 MAOWNLOQOLDTRYLKQHLQLYSDTFPMELRQFLAPWIESQDMAYAAKESHATLVFHNL 60
        |||||||
Db      1 maownloqoldtryleqhlqlysdtfpmeleqlapwiesqdmayaaakeshatlvfhn 60
OY      61 LGEIDQOYSRELIQESNVLYQHNLRLIKOFLORSYLEKPMELIRIYARCLMESSRLLOTA 120
        |||||||
Db      61 lgeidqoysreliqesnvlyqhnrlrikoflorsylekpmeliriyarclmesselotaa 120
OY      121 TAAOOGCGQNHFFTAAVVEREKOQMLEOHLODVRKRYODLEOKKKVVENLODDDFNYKTLK 180
        |||||||
Db      121 taaoogcgqnhfftaavverekomleohlodvrkrvodyleokkkvvenlodddfnyktlk 180
OY      181 SOGDMODLGNNGSVYTRKQMQL EOMLTALDOMRNS IYSELAGLLSMEYVOKTLTDEEL 240
        |||||||
Db      181 sgdmmdlgnngsvytrkqmql eomltaldomrns iyseलगल्लमेयवक्ल्लिदेे 240
OY      241 ADMKRRPEIACIGSPNICLDRLENWITSLAESOLOTROQIKLELOOKVYKGDPIVQ 300
        |||||||
Db      241 admkrrpeiacigspnicldrlewnitslaesolotroqiklelookvyskgdptvq 300
OY      301 HRPMLEERIVLELFRNLKSAFVVERQPCMPMHPDRPLVTKGVQFTTKVRLVKPEELNY 360
        |||||||
Db      301 hrpml eerivlelfrnlksafvverqpcmpmhpdrplvltkqvfttkvrlvkkpeelny 360
OY      361 QLKIKVCTDKDGDVAALALGRKRFNLCNTYKVMNNEESNNGSLAEFKHLTLRQRCGN 420
        |||||||
Db      361 qlkikvcidkdsqdvaaalrgsrkfnlglntkvmmeesnglsjaefkhlilreqrcgn 420
OY      421 GGRANCDASLIYTEELHLITFETEVYHOGKIDLETHSLPVAVINSICOMPANASILMY 480
        |||||||
Db      421 ggrancdasliyteelhlitfetevyhgkikidlethslpvvniicompawasilwy 480
OY      481 NMLTNPKNVNFFTKPRPGTMDVAEVLISWOFSSSTTKRGLSIEQLTTLAEKLLGCVNYS 540
        |||||||
Db      481 nmltnpkvnfftkprpgtmdvaevliswofsssttkrglsieqlttlaekllpgvnyys 540
OY      541 GCOITWAKFCENMAGKGSFWVWLDNIIDLKYYILALMNGGYIMGFISKREERAILST 600
        |||||||
Db      541 gcoitwakfckenmagkgsfwwldnidlvkkyilalwmegylimgfiskerailstc 600
OY      601 KPPGFTLLRFSSSKEGVFTWVERKDISGKTQIQSVEPYTRKOOLNNMSFAEIMGYKIM 660
        |||||||
Db      601 kppgftllrfssskegyftfwvekdlsqkctqigsvepytkqqlnmsfaeilmgylim 660
OY      661 DATNIIIVSPVLYLPDIPKEEAFGKYCRPESQHPHADGSAAPYIKTKFICVPTTCSN 720
        |||||||
Db      661 datniiivsplylvpdipkeeafigkycrpesqhpdaadpysaapyiktkficvptlcsn 720
OY      721 TIDLPMSPRTLDLSLMQFCNNGCAEPSAGQFESLTFDMDLTSECATSPM 770
        |||||||
Db      721 tidlpmsprtldslmqfngngeaepsagqfesltfdmdltsecatspm 770

```

```

RESULT 5
AAV03768
ID      AAV03768 standard: Protein: 770 AA.
XX      AAV03768:
AC      11-JUN-1999 (first entry)
DT      Human STAT3 allelic variant.
XX      DE      Signal Transducer and Activator of Transcription 3; STAT3; allele; IL-6;
XX      intracellular transcription factor; interleukin-6; medicament; variant;
KW

```

KW pharmaceutical; autoimmune disease; inflammatory; human.
 XX
 OS Homo sapiens.
 PN EP905234-A2.
 XX
 PD 31-MAR-1999.
 XX
 PF 18-FEB-1998; 98EP-0102774.
 XX
 PR 16-SEP-1997; 97EP-0116061.
 XX
 PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
 PI Della Pietra L, Serlupi-Crescenzi O;
 DR MPI: 1999-192664/17.
 XX N-PSDB; AAX29281.
 PT New human signal transducer and activator of transcription 3 (STAT3)
 PT allelic variant useful for treatment of autoimmune and inflammatory
 PT disease
 XX
 PS Claim 2; Page 9-13; 32pp; English.
 XX
 CC The present sequence represents a predominant allelic variant of human
 CC signal transducer and activator of transcription 3 (STAT3) protein, an
 CC intracellular transcription factor which mediates IL-6 signals. The
 CC encoding sequence differs from the original published human STAT3 gene
 CC sequence (Akira et al., 1994). Host cells comprising a recombinant STAT3
 CC DNA molecule can be used for the recombinant expression of the variant.
 CC STAT3 protein is useful as a medicament or pharmaceutical composition for
 CC treatment of autoimmune or inflammatory diseases.
 CC
 SQ Sequence 770 AA:
 Query Match 99.4%; Score 4006; DB 20; Length 770;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 765; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 OY 1 MAQNDLQDLDTRYLKHLYSDTPMELROFLAPMIESODMAYAAKSHATLVFHNH 60
 DB 1 maqndlqgdtdrylqhlqlydsfpmelrqlapiesgdwayaaskshalvfhnl 60
 OY 61 LGEIDQYSFLEDSNVLYQHNLRIKQFLQSRYLEKPMELIARIVARCLMEESRLQTA 120
 DB 61 lgeidqysrflqesnvlyqhnlrrikqlqsfrylekpmelariarclweesrlqtaa 120
 OY 121 TAAQGGQANHPRAAVVTEKQMLEQHLQDVKKRVODLEQKKAVVEMLODDPFPNRYTLK 180
 DB 121 taaqggqanhprraaavvteqqlqhlqdvkrkvodleqkkavvemloddpfpnrytlk 180
 OY 181 SOGDMDLNGNNSVTRKQKQLEQMLTALDQKRSTVSLAGLSAMEYVOKLTDEEL 240
 DB 181 sqgdmldngnnsvtrkqkqlqemltaldqkrrstivslaglsameyvkltdeel 240
 OY 241 ADMKRPETLACTIGCPNLCIDLENNITSLAESQLQTRQOIKKLEELQKVSKGDPYV 300
 DB 241 admkrrpetlactigcpnlicidlenntslaesqlqtrqoikkleelqkvskgdpiv 300
 OY 301 HRPMLEERIVELFRNLKMSFVVERPCMPHMRDRPLVITGVQFTTKVNLVYKPEPLNY 360
 DB 301 hrpml eerivel frnlkmsf vverpcmp hmrdrplv itgvqfttkvnlv kpeplny 360
 OY 361 QLKIKVICDKSGDVAALGSRKFNILGTNTKVMNEESNGSLSEFKLTLREQSCGN 420
 DB 361 qlkikvicdksgdvaalgsrkfniltgntkvmneesngslsefkltlreqscgn 420
 OY 421 GGRANDASLVTTELHLTFETEVYHQGLKIDLETHSLPVVVISNIGCPNMAASILMY 480
 DB 421 ggrandaslvttehlhtfetevyhgkldlethslpvvvisnigcpnmasilmy 480

OY 481 NMLTNNPKNVNEFTKPPIGTWQVAEVLWSQFSSTTKRGLSIEQLTTLAEKLGPGVNS 540
 DB 481 nmltnnpk nvneftkppigt wqvaelw sqfss tkrgl sieql t tlaeklgpgvns 540
 OY 541 GCQITWMAFKCKENMAGKGFSPVWLDNIIIDLVKKTIIALMNEGYIMGTISEREALLST 600
 DB 541 gcqitwmafkck enmagk gfs p v w l d n i i d l v k k t i i a l m n e g y i m g t i s e r e a l l s t 600
 OY 601 KPPGTELLRFSESSKEGAVTFWWEKDISGKTQIOSVPPYTKQOOLNNNSPFEIIMGYKIM 660
 DB 601 kppgte llrfse sskegav t f w w e k d i s g k t q i o s v p p y t k q o o l n n n s p f e i i m g y k i m 660
 OY 661 DATNLSPIVLYLPDIPEKEAFCKRYCRPESQHPHADPGSAAPYLKTKFTICVPTTCNSN 720
 DB 661 datnls p i v l y l p d i p e k e a f c k r y c r p e s q h p h a d p g s a a p y l k t k f t i c v p t t c n s n 720
 OY 721 TIDLPMSPTLDSLMQGNNGEGAPPSAGGFESLTFPMDLTSECATSPM 770
 DB 721 tidlpmspt lds lmqgnng eap s aggfes ltfpmd ltsecat spm 770
 RESULT 6
 AAB58442
 ID AAB58442 standard; Protein: 793 AA.
 XX
 AC AAB58442;
 XX
 DT 14-MAR-2001 (first entry)
 XX
 DE Lung cancer associated polypeptide sequence SEQ ID 780.
 XX
 KW Human; lung cancer associated protein; neuroprotective; cytoskeletal;
 KW gastrointestinal; nephrotropic; antineoplastic; immunomodulatory;
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
 KW proliferative disorder; wound healing; infectious disease.
 XX
 OS Homo sapiens.
 OS
 PN W020005180-A2.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05918.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GRNOME SCI INC.
 PA (ROSE/) ROSEN C A.
 XX
 PI Ruben SM;
 PI
 DR WPI: 2000-587514/55.
 DR N-PSDB; AAF18318.
 PT Lung cancer associated gene sequences, referred to as lung cancer
 PT antigens, useful for treatment, prevention, and diagnosis of disorders
 PT such as lung cancer -
 XX
 PS Claim 11; Page 1310-1313; 1425pp; English.
 XX
 CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
 CC associated proteins and polynucleotide sequences, their agonists, and
 CC antagonists may have neuroprotective, cytoskeletal, cardioactive;
 CC immunomodulatory; muscular active general; vlnnerary; gastrointestinal
 CC general; nephrotropic; antineoplastic; gynecological; or antibacterial
 CC activity. The invention also includes antibodies specific for the
 CC protein or polynucleotide sequences. The lung cancer associated
 CC polynucleotide sequences may be used for detection of lung cancer,
 CC chromosome identification, as chromosome markers, and for numerous other
 CC diagnostic or research purposes. The proteins may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,

CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative disorders. The proteins may also be used in the treatment of wounds and infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and CC peptide AAB58549 are used in the course of the invention for the CC identification and characterisation of the polynucleotide and protein sequences.

XX
XX
SQ Sequence 793 AA:

Query Match 99.4% Score 4006: DB 21: Length 793:
Best Local Similarity 99.4% Pred. No. 0:
Matches 765: Conservative 4: Mismatches 1: Indels 0: Gaps 0:

1 MAQNNLOQDTRYLKOLHQLYSDTFPMELRQFLAPMIESODMAAASKESHATLVFHNL 60
24 magwnglqldtrylqeqhqlgysdfmelrtqlapwiesgdwayaaskeshatlvfhn 83
61 LGETDOQYSRFLOSNNVLYOHNLRIKQFLOSRYLEKPMELARIYARCLMEESRLQTA 120
84 lgeidqysrflgesnvlqynhlrlrlkqlqysrlyekpmetarivarcilweestrllqta 143
121 TAAOOGGGAHPHTAAVYTEKQOMLEOHLDVYRKRVQDLEQMKVENVLDQDFNYKTLK 180
144 taaggggqanphtaaavvtekgmqleqhlqdvcrkvqdlqekmkvvenlqddfdnkytkl 203
181 SQGMODLNGNOSVTRKQKQMLEQMLTALDOMRSTVSELAGLLSMEYQKTLTDEEL 240
204 sqgdmqdnngnsvtrkmgqleqmltalqdmrsvselagllsmevqkltldecl 263
241 ADMKRREPIACIGPPNICDLRLNNMITSLSAESOLTRQOKKLEELQOKSYKGDPIVQ 300
264 admkrrtgiacigppnicdlrlennwitslsaesqlqtrqkklleelqkksygdpiqv 323
301 HRPMLEEVIVLFRLNLSAFVVERQPCMPHPRPLVTKGVQFTTKVRLVKFPELVN 360
324 hrpmleevivlfrlnlsafvverqpcmpmhprplvtkgvqfttkvrlvkfpelevn 383
361 QLKIKVCIDKSDGVAALRGRKFNILGTNTKVMNMEESNGSLSAEKKLTLREORCGN 420
384 qlkikvcidsdgvaaalrgsrkfnilgtntkvmnmeesngslsaekhltilregrcn 443
421 GGRANCDASLIVTEELHILFETEVYHOGKIDLETSHSLPVVVISNICOMPMAASITWY 480
444 ggrancdaslivteelhilfetevyhnglkidletshslpvvvisnicompmaasitwy 503
481 NMLTNNPNVNFETKPPICGTMDQVAEVLVSMQSSSTRKGLSTEDLTTLAEKLLGPNVYS 540
504 nmltnnpnvnfetkppicgtmdovaevlsvmqssstrkglsteditltaekllgpnvys 563
541 GCQITWAKFCKENNAAGKGFSPWMLNDIIDLVKRYIALMNEGYIMGFISKEERAILST 600
564 gcqitwakfckennagkgsfswvldmldivkkyialalmnegyimgfiskerailst 623
601 KRPCTELLRFESSEKREGVFTTWYKDISGKTQIQSVFPTKQOLNNKSPFELIMGKIM 660
624 krpctellrfesssekregvfttwykdsgktqiqsvfptkqolnnkspfelimgkim 683
661 DATNIVLSPLVLYLPDIPKEEAFKCYCRPESQHPHEADPGSAAYLYLKTFTICVPTTCSN 720
684 datnivlsplvlylpdipkeefkcycrpesqhpheadpgsaaylylktfticvpttcsn 743
721 TIDLPMSRITLDSLMQFNGNKGAEPSAGGOFESLTFDMDLTSECANSPM 770
744 tidlpmsriltldslmqfngnkgaeapsaggofesltfdmdltsecanspm 793

RESULT 7
AAB19964 ID AAB19964 standard: Protein: 770 AA.
XX
AC AAB19964:

DT 28-MAR-2001 (first entry)
XX Human signal transducer and activator of transcription STAT-3.
DE
XX
KM STAT-3; signal transducer and activator of transcription 3;
XX human; crystal; drug screening.
XX
OS Homo sapiens.

XX Key Location/Qualifiers
FH Domain 134..320
FT /note="coiled-coil domain"
FT 321..493
FT /note="DNA-binding domain"
FT Domain 494..583
FT /note="linker domain"
FT Domain 584..688
FT /note="SH2 domain"
FT Region 689..717
FT /note="C-terminal tail segment"
FT Modified-site 705
FT /note="O-phosphorylated"

XX US6160092-A.
XX 12-DEC-2000.
XX
XX 29-MAY-1998: 98US-0087465.
XX
XX 29-MAY-1998: 98US-0087465.
XX
XX (UYRO) UNIV ROCKEFELLER.

XX Chen X, Darnell JE, Kuriyan J, Vinkemeier U, Zhao Y, Jeruzalmi D;
XX WPI: 2001-101568/11.
XX DR N-PSDB: AAA89229.
XX

XX Novel crystal useful in drug screening assays, comprises portion of
PT signal transducer, activator of transcription and duplex DNA -
XX
XX
PS Disclosure: Column 67-71: 206pp: English.

XX The present sequence is that of human signal transducer and
CC activator of transcription 3 (STAT-3). The invention provides a
CC crystal of a core portion of a STAT protein in dimer form with an
CC 18-mer duplex DNA (see AAB89233) that contains a binding site for the
CC STAT dimer. The core portion comprises a coiled-coil domain
CC comprising 4 long helices, a DNA binding domain which contains an
CC immunoglobulin-like fold, a C-terminal SH2 domain and a domain that
CC links the DNA binding and SH2 domains. The crystal is of sufficient
CC quality to perform X-ray crystallography studies. Methods of
CC preparing the crystals are included in the invention. Knowledge of
CC the STAT protein's 3-dimensional structure will aid in structure-based
CC drug design. The crystal can be used in drug screening assays to
CC identify agonist and antagonist compounds. Antagonists can be used
CC to treat inflammation, allergy, asthma and leukaemia, and agonists
CC viral diseases, growth retardation, thrombocytopenia, cancer, obesity,
CC characterized by insufficient STAT activity.

XX
SQ Sequence 770 AA:

Query Match 98.7% Score 3977: DB 22: Length 770:
Best Local Similarity 98.6% Pred. No. 0:
Matches 759: Conservative 6: Mismatches 5: Indels 0: Gaps 0:

OY 1 MAQNNLOQDTRYLKOLHQLYSDTFPMELRQFLAPMIESODMAAASKESHATLVFHNL 60
Db 1 magwnglqldtrylqeqhqlgysdfmelrtqlapwiesgdwayaaskeshatlvfhn 60
OY 61 LGETDOQYSRFLOSNNVLYOHNLRIKQFLOSRYLEKPMELARIYARCLMEESRLQTA 120

|||||
Db 1geidqgsrflqgsnvlyqhnllrrlxlqfllqstyllekpmelariavarcilweesrlllqltaa 120
OY 121 TAAOOGGANHPRTAAVYTEKOOEOLHODVRRKRVODEOKMKVVENLODDDFNYKTLK 180
Db 121 taaqgqganhpptaavvtekgmleqnhqdvtrkvrqdlleqkmkvenlqdddfnykcllk 180
OY 181 SOGDMODLNGNOSTROKMOOLEOMLTALDOMRRSISVSELAGLLSAMEYOKTLTDEEL 240
Db 181 sqgdmqdingnqsvtrckmqqlleqmltalqdmrrisvselagllsameyqkltldeei 240
OY 241 ADMKRRPEIACIGRPNICDLRLEMMITSLAESOLQTRQOIKKLELOOKSYKSDPIVQ 300
Db 241 admkrrpqiaclgppnclclrlenwltslaesqltrqqlkkleelhqkvsykgdplvq 300
OY 301 HRPMLERIVELFRNLMSAFVVERQPCMPHDPRLVIKTCVQPTTKVRLLVKPEELNY 360
Db 301 hrpmlerivelfrnlmsafverqpcmpmhdpdrplviktcvqpttkvrlvlvkfpeiny 360
OY 361 QLKIKVCIDKDSGVAALRGSRKFNIIGTNTKVMNMESNNCSLSAEFKHLTLRBORCN 420
Db 361 qlkikvcidkdsqdvaaalrgsrkfnllgtntckvmmeesngslsaefkhltlrqrqcn 420
OY 421 GGRANCA SLVTEELHLITFETEVYHOGKIDLETHSLPVVISINICOMPNAASILMY 480
Db 421 ggrancaslvteelhilitfetevyhgikidlethslsvvisnlcqmnpawasilly 480
OY 481 NMLTNPNKVNVEFTKPIGTDQVAEVLWSQSSPTTKRGLSIEQJTTLAEKLLGCVNYS 540
Db 481 nmltnpnkvnvftkppigtqdwqvaevlswqsfstkrqlsleqjttlaeklllpgvnyvs 540
OY 541 GCOITWAKFCKENMAGKGFSEFWMLDNIIDLVKXYIALAMEGYIMGFTISKERERAILST 600
Db 541 gcqitwanfckemagkgsfywvwlidnildlvkxyialamegyimgfiskereraillst 600
OY 601 KPPGTFILRFESSESKEGVTFTWWEKOLISGKTQIOSVEPTYTKOOLNNNSFAELIMGYKIM 660
Db 601 kppgtfllrfesssksgvttfwekdlsyktlqsvpytkqlnnmsfaelimgykim 660
OY 661 DATNIIIVSLVLYLPDIPKEEAFGKYCRPESOEHPHEADPGSAAPYLKTKFCVPTTCSN 720
Db 661 datniiivslvlylpdipkeeafigkycrpesgehpheadpgsaapylkktfcivpttcsn 720
OY 721 TIDLPMSPRTLDLMQFGNNGEAPSPAGGOFESLTFPMDLTSSECATSPM 770
Db 721 tidlpmspiraldslmqfgnngaeapsagqfeslftdmeiltsecatspm 770
RESULT 8
AAR82993
ID AAR82993 standard; protein: 770 AA.
XX
AAR82993:
AC
XX
DT 25-MAR-1996 (first entry)
XX
DE Human placenta acute phase response factor protein.
XX
KW human; acute phase response factor; transcription factor;
KW interleukin-6; signal transmission; placenta; antibody; antise; se;
KW ribozyme; antiinflammatory; antitumor; hypotensive; therapy.
XX
OS Homo sapiens.
XX
PN EP676469-A2.
XX
PD 11-OCT-1995.
XX
PF 29-MAR-1995; 95SEP-0104670.
XX
PR 04-APR-1994; 94JP-0065825.
XX
PA (KISH/) KISHIMOTO T.

XX Akira S, Kishimoto T;
PI
XX
DR WPI: 1995-346089/45.
DR N-PSDB: AAT05616.
XX
XX New acute phase response factor - for developing inhibitory agents
PT for treating diseases induced by cytokine(s) such as IL-6, e.g.
PT inflammatory diseases
XX
PS Claim 3; Page 9-12; 31pp; English.
XX
CC The sequence corresponds to a human acute phase response factor
CC (APRF), a transcription factor related to signal transmission of
CC Interleukin-6 (IL-6). The protein is expressed from a human
CC placenta cDNA, isolated using an IL-6-treated mouse liver cDNA
CC probe. APRF-inhibitors, e.g. antibodies, antisense
CC oligonucleotides or ribozymes, may be used to treat diseases
CC induced by IL-6, e.g. inflammatory disease, leukemia, cancer,
CC osteoclastia, pulmonary hypertension, etc.
XX
SQ Sequence 770 AA:
Query Match 98.5%; Score 3970; DB 16; length 770;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 758; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
OY 1 MAQNWQLOQDLDRYLKQHLQSLDFFPMELRQFLAPWIESQWMAAASKESHATLVFNL 60
Db 1 maqnwqlqldrllyleqlhqlqysdsfpmelrqlapwiesqdwyaaskeshatlvfnl 60
OY 61 IGEIDQOVSRLFQESNVLYQOHLRRIKQFLOSRYLEKPMELARIYARCLMEESRLQTTAA 120
Db 61 lgei lqsrllqgsnvlyqnhllrrlxlqfllqstyllekpmelariavarcilweesrlllqltaa 120
OY 121 TAAOOGGANHPRTAAVYTEKOOEOLHODVRRKRVODEOKMKVVENLODDDFNYKTLK 180
Db 121 taaqgqganhpptaavvtekgmleqnhqdvtrkvrqdlleqkmkvenlqdddfnykcllk 180
OY 181 SOGDMODLNGNOSTROKMOOLEOMLTALDOMRRSISVSELAGLLSAMEYOKTLTDEEL 240
Db 181 sqgdmqdingnqsvtrckmqqlleqmltalqdmrrisvselagllsameyqkltldeei 240
OY 241 ADMKRRPEIACIGRPNICDLRLEMMITSLAESOLQTRQOIKKLELOOKSYKSDPIVQ 300
Db 241 admkrrpqiaclgppnclclrlenwltslaesqltrqqlkkleelhqkvsykgdplvq 300
OY 301 HRPMLERIVELFRNLMSAFVVERQPCMPHDPRLVIKTCVQPTTKVRLLVKPEELNY 360
Db 301 hrpmlerivelfrnlmsafverqpcmpmhdpdrplviktcvqpttkvrlvlvkfpeiny 360
OY 361 QLKIKVCIDKDSGVAALRGSRKFNIIGTNTKVMNMESNNCSLSAEFKHLTLRBORCN 420
Db 361 qlkikvcidkdsqdvaaalrgsrkfnllgtntckvmmeesngslsaefkhltlrqrqcn 420
OY 421 GGRANCA SLVTEELHLITFETEVYHOGKIDLETHSLPVVISINICOMPNAASILMY 480
Db 421 ggrancaslvteelhilitfetevyhgikidlethslsvvisnlcqmnpawasilly 480
OY 481 NMLTNPNKVNVEFTKPIGTDQVAEVLWSQSSPTTKRGLSIEQJTTLAEKLLGCVNYS 540
Db 481 nmltnpnkvnvftkppigtqdwqvaevlswqsfstkrqlsleqjttlaeklllpgvnyvs 540
OY 541 GCOITWAKFCKENMAGKGFSEFWMLDNIIDLVKXYIALAMEGYIMGFTISKERERAILST 600
Db 541 gcqitwanfckemagkgsfywvwlidnildlvkxyialamegyimgfiskereraillst 600
OY 601 KPPGTFILRFESSESKEGVTFTWWEKOLISGKTQIOSVEPTYTKOOLNNNSFAELIMGYKIM 660
Db 601 kppgtfllrfesssksgvttfwekdlsyktlqsvpytkqlnnmsfaelimgykim 660
OY 661 DATNIIIVSLVLYLPDIPKEEAFGKYCRPESOEHPHEADPGSAAPYLKTKFCVPTTCSN 720

```

Db 661 datnlllplylyltpdipkeeaifgkyrcpesgehpdeadpgsaapytktkficvptlcsn 720
Oy 721 TTDLPMSPTLDSLMOPGNNGEAPSGAGCPESLTFDMDLTSECATSPM 770
Db 721 tldlpmSpraldslmqfmggeapsgagqflesltfmdltsecatspm 770

RESULT 9
ID AAY72842 standard; protein: 393 AA.
AC AAY72842;
DT 31-MAY-2001 (first entry)
DE Mouse Stat3 protein fragment #3 (378-770 amino acids).
KM Mouse; Stat3 protein: transcription factor; c-Jun; gene transcription;
cellular transformation; dysproliferative disease; cancer; psoriasis;
therapy.
KW Mus musculus.
OS WO200116605-AZ.
PN 08-MAR-2001.
PF 30-AUG-2000; 2000WO-US23822.
PR 31-AUG-1999; 99US-0387418.
PP (UYRO ) UNIV ROCKEFELLER.
PI Zhang X, Horvath C, Wrzeszczynska MH, Darnell JE;
DR WPI: 2001-226705/23.
XX Identifying an agent for use in modulating the interaction between
transcription factor c-Jun and a Stat3 protein -
XX Example 2; Page 68-69; 86pp; English.
TS The present sequence is mouse Stat3 protein fragment containing 378-770
XX amino acids of Stat3 protein. This Stat3 fragment showed very weak
XX binding to c-Jun protein in the cell extract.
XX The invention relates to methods for identifying interacting regions of
XX transcription factors and methods for identifying agents which modulates
XX the interaction between a transcription factor such as c-Jun and a Stat
XX protein such as Stat-1 and Stat-3, useful for modulating gene
XX transcription e.g., cellular transformation. These identifying agents are
XX used in the treatment of dysproliferative diseases and also for treating
XX cancer and psoriasis. A Stat protein comprises the N-terminal domain,
XX coiled-coil domain, DNA binding domain, linker domain, SH2 domain and
XX transactivation domain.
SQ Sequence 393 AA:

Query Match 51.8%; Score 2087; DB 22; Length 393;
Best Local Similarity 100.0%; Pred. No. 3.5e-172;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 378 LRSGRRFNILGNTKVMNNEESNNGSLSAEFKHLTLREORCGNGRANCDA SLVTEELH 437
Db 1 lrgsrkfnllgntkvmmesngslsae fkhltlreqrcngngrandaslvtteeh 60
Oy 438 LITFTEVYHQSGLIDLEHSLPVVVISNICQPMNMAASILMYNMLTNPKNVNFTKPP 497
Db 61 lltfetevyhgsjklidlehs lppvvishncqpmnasilw ynm ltnpknvnfttkpp 120
Oy 498 IGTQWQVAEVLSSQFSTTKRGSLISQLTLEKLLGPGVNVSGCQITWAKCKCKEMAGK 557
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

```

```

Db 121 lgtwdyvaej]swgfssltkrgisicqltllaekllgpyvnysgqjtwakickenmagk 180
Oy 558 GFSEFWALDNIIDLVKKYILALWNEGYIMGFISKERERALLSRPGTFLLRFSSESKG 617
Db 181 gsfewwldnidlvkkyilalwnegyimgfiskereral]slkpgtcllirfseskeg 240
Oy 618 GVTFTWEXDKISGTOIOSVEPYTKOQLNNMSPAEIIMGYKIMDATNILLVSLVYLYPUI 677
Db 241 gvtftwexdkisgtioiosvepytkqqlnmsfaeelimgyk imdatnillvslvylypdi 300
Oy 678 PKEEAFGKYCRPESGHPADPGSAAPYLKTKFICVPTTCSTNTIDLPMSPRTLUSIMOF 737
Db 301 pkeeaifgkyrcpesgehpdeadpgsaapytktkficvpttcstntidlpmSpraldslmqf 360
Oy 738 GNGEGAPSGAGCPESLTFDMDLTSECATSPM 770
Db 361 gnngeapsgagqflesltfmdltsecatspm 393

RESULT 10
ID AAR72078 standard; protein: 750 AA.
AC AAR72078;
DT 27-SEP-1995 (first entry)
DE Human Stat91.
XX Signal transducer and activator of transcription; ISGF-3; STAT;
XX Stat3; receptor recognition factor; transcription factor;
XX cellular debilitation; derangement; dysfunction; interferon-alpha;
XX interferon-gamma.
XX Homo sapiens.
XX OS WO9508629-A.
XX 30-MAR-1995.
XX 26-SEP-1994; 94WO-US10849.
XX 24-SEP-1993; 93US-0126588.
XX 24-SEP-1993; 93US-0126595.
XX 11-MAR-1994; 94US-0212184.
XX 11-MAR-1994; 94US-0212185.
XX (UYRO ) UNIV ROCKEFELLER.
PI Darnell JE, Schindler CW, Shuai K, Wen Z, Zhong Z;
DR WPI: 1995-139598/18.
XX N-PSDB; AA089335.
XX Receptor recognition factor implicated in transcriptional
XX stimulation of genes - useful in drug screening assays and/or
XX for treating cellular debilitations, derangements and/or
XX dysfunctions, etc.
XX Disclosure: Page 84-88; 160pp; English.
XX The sequences of cDNA encoding receptor recognition factors having
XX mol.wt. of 113 kDa (Stat113), 91 kDa (Stat92) and 84 kDa (Stat84)
XX are given in AA089335-37 and the deduced amino acid sequences of the
XX STAT proteins in AAR72077-79. These ISGF-3-derived proteins are
XX activated by binding of interferon-alpha (all 3 Stat proteins) or
XX interferon-gamma (Stat91) to cell receptors.
SQ Sequence 750 AA:

Query Match 50.1%; Score 2017; DB 16; Length 750;
Best Local Similarity 52.5%; Pred. No. 1.2e-165;

```

Matches 389; Conservative 145; Mismatches 193; Indels 14; Gaps 9;				
OY	1	MAOMNLOOLDTRYLKQHLQLYSDTFPMELOFLAPWIESODMAYAAKSKESHATLVFHNL	60	
Db	1	msqwyelqgidskrlqyhyldsfpmeltrqylaqvlekq*ehaandvsfatirfndl	60	
OY	61	LGELIDQOYSRFLQESNVLYOHNLRIRIKOFLQSRYLEKPMELIARIVARCLMEESRLQTA	120	
Db	61	lsqlddysrflslenfllqhnirskrnldqndfndpigmmllyscskeerklleaq	120	
OY	121	TAADQGGQANHPTAAVYTEKQOMLEOHLODYRKRVQDLEQKMKVVENQDDPFRNYTKL	180	
Db	121	rftqg--aqsgnlgstvmldkqkeldskvrvnkdymcshelkstdedqdeydfcktl-	177	
OY	181	SOGMODLNGNOSVTRQKMOOLEOMLTALDOMRSIYSELGLISAMEYVQKTLTDEEL	240	
Db	178	-gnrehetnyaksdqkqeglllkmylmldnkrkevnhkltellnvtelqnalindel	236	
OY	241	ADMKRREIACIGGPRNICDLRENMTSLAESQLQTRQOIKLEELQOKVSYGDPRIQ	300	
Db	237	vwktrqgsacigppnacdqlqmwftivaeslqgvtrqklkleeqkytyehdptlk	296	
OY	301	HRPMLEERIVLEFRNLKMSAFVVERQPCMPHNDPRPLVITKGVOFTTRVRLVFRPELNY	360	
Db	297	nkqylwdrtfslfqglqgsstfverqpcmrphqprplvltkgvftvklrlvlkqeln	356	
OY	361	OLKIKVCIDKSDGVAAALRGSRKFNILQTNTKVMNMESNNGSLSAERKHLTLREQRCN	420	
Db	357	nlkxkvldfckvnerntvkgfrfnlqgthckmmesstngslaaefrlqklegk--n	414	
OY	421	GGRANDASLIVTEELHLITETEYVHOGKIDLETSHLPVVVISNICOMPANASILIWY	480	
Db	415	agrtneqplivteelshstfctqcpqglvldletslslyvvlsvnsqglpsyaasllwy	474	
OY	481	NMLNNKKNVNEFFRKPGTMDQVAYEVLSKQFSSSTRKGLSIBOLTTLAEKLLGPGVYS	540	
Db	475	nmlyaeprnslfllprraqaqlsevlswqfsvtkqglvndqimgelllqpnaspd	534	
OY	541	GCQTWAKFCCKENMAGKGFPMWLDNIDLVKKYIALNMEGYIMGPISTKEREALIST	600	
Db	535	gl-lpwtrfckenindkpfwlvieslletlklhlpwdngcimgfiskereralld	593	
OY	601	KPRGTFLLRESSEKQGVTFTWVEKDIS-GKTQIQSVEPTKQOOLNMKSAEILIMGKI	659	
Db	594	qgpgtfillrfsessregaltfllwersqnggepdflhavepytlkelsavtfdlrmkyk	653	
OY	660	MDANILIVSPLYVLPDIPEKEAFGK-YGRP-ESOEHPDAPGSAAPYLKTKFKICVT---	714	
Db	654	maeenlpenplkylpnlidkdhafgkystrpkeapemeldgprkgtgylkltellsvsevh	713	
OY	715	PTTCSNTID-LPMSPRITLDSL 734		
Db	714	psrlqtdtlpmspeefdev 734		
RESULT 11				
AAW03168				
ID AAW03168 standard; Protein; 750 AA.				
AC AAW03168:				
XX 24-OCT-1996 (first entry)				
DT Human STAT1-alpha.				
XX STAT1-alpha; signal transducer and activator of transcription;				
KM DNA binding protein; ligand: receptor; oncogenesis; inflammation;				
KM autoimmune disease; antagonist; therapy; STAT1.				
XX Homo sapiens.				
XX Key Location/Qualifiers				
FH Domain 399..508				

FT	/label= DNA-binding-domain
FT	/note= Claim 3, page 110*
FT	Modified-site
FT	701
FT	/label= Phosphorylation_site
FT	Modified-site
FT	727
FT	/label= Phosphorylation-site
XX	
PN	MO9620954-A2.
XX	
PD	11-JUL-1996.
XX	
PF	28-DEC-1995: 95MO-US17025.
XX	
PR	06-JAN-1995: 95US-0369796.
XX	
PA	(UYRQ) UNIV ROCKEFELLER.
XX	
PI	Darnell JF, Horvath CM, Wen Z, Zhong Z;
XX	
DR	WPI: 1996-333941/33.
XX	
DR	N-PSDB: AAT 476.
XX	
PT	New STAT protein DNA-binding domain peptide(s) - useful for
PT	diagnosing, preventing or treating cellular dysfunction, e.g.
PT	oncogenesis, inflammation, parasitic disease or autoimmunity
XX	
PS	Disclosure: Page 69-73; 138pp; English.
XX	
CC	Signal transducer and activator of transcription (STAT) protein
CC	STAT1-alpha (AAW03168), also known as STAT1, is a 91 kDa protein
CC	having a dual purpose, i.e. signal transduction from ligand-activated
CC	receptor kinase complexes followed by nuclear translocation and DNA
CC	binding to activate transcription. Recombinant STAT1-alpha can
CC	be obtd. using an isolated cDNA clone (AAT31276). STAT1-alpha
CC	includes a DNA-binding domain (see also AAW03165) capable of both
CC	receptor recognition and message delivery via DNA binding in a
CC	receptor-ligand specific manner. Maximum STAT1-alpha activation
CC	of genes requires phosphorylation of both Tyr-701 and Ser-727.
CC	STAT proteins and their DNA binding domains (see also AAW03165-67,
CC	AAW03169-76) are useful for screening antagonists used to inhibit
CC	STAT-mediated signal transduction and activation of transcription.
XX	
SO	Sequence 750 AA:
Query Match	
Best Local Similarity 52.5%; Pred. No. 1,2e-165;	
Matches 389; Conservative 145; Mismatches 193; Indels 14; Gaps 9;	
OY	1 MAOMNLOOLDTRYLKQHLQLYSDTFPMELOFLAPWIESODMAYAAKSKESHATLVFHNL
Db	1 msqwyelqgidskrlqyhyldsfpmeltrqylaqvlekq*ehaandvsfatirfndl
OY	61 LGELIDQOYSRFLQESNVLYOHNLRIRIKOFLQSRYLEKPMELIARIVARCLMEESRLQTA
Db	61 lsqlddysrflslenfllqhnirskrnldqndfndpigmmllyscskeerklleaq
OY	121 TAADQGGQANHPTAAVYTEKQOMLEOHLODYRKRVQDLEQKMKVVENQDDPFRNYTKL
Db	121 rftqg--aqsgnlgstvmldkqkeldskvrvnkdymcshelkstdedqdeydfcktl-
OY	181 SOGMODLNGNOSVTRQKMOOLEOMLTALDOMRSIYSELGLISAMEYVQKTLTDEEL
Db	178 -gnrehetnyaksdqkqeglllkmylmldnkrkevnhkltellnvtelqnalindel
OY	241 ADMKRREIACIGGPRNICDLRENMTSLAESQLQTRQOIKLEELQOKVSYGDPRIQ
Db	237 vwkttrqgsacigppnacdqlqmwftivaeslqgvtrqklkleeqkytyehdptlk
OY	301 HRPMLEERIVLEFRNLKMSAFVVERQPCMPHNDPRPLVITKGVOFTTRVRLVFRPELNY
Db	297 nkqylwdrtfslfqglqgsstfverqpcmrphqprplvltkgvftvklrlvlkqeln

PT obtaining products for treating e.g. inflammation, allergy, asthma,
 PT anaemia, cancer, obesity, viral disease and growth retardation
 XX
 XX
 PS Claim 20: Page 77-79: 11pp: English.
 XX
 CC This is human Stat1-alpha protein. The invention describes method
 CC of producing milligram quantities of 3 forms of purified Stat1
 CC protein: Stat1-alpha and Stat1-beta (see AAM62995) from baculovirus
 CC infected insect cells, and Stat1c (see AAM62996) from E. coli.
 CC Stat1-alpha has an N-terminal compact domain that enhances the
 CC DNA binding of the Stat protein. The Tyr-701 residue is
 CC phosphorylated in vivo. In vitro phosphorylation of the protein
 CC plus cysteine residue alkylation to prevent aggregation, coupled to
 CC a chromatography protocol, has allowed the purification of
 CC activated Stat proteins. The Stat proteins and fragments can be
 CC used to identify antagonists and agonists of Stat function.
 CC Antagonists can be used to treat e.g. inflammation, allergy,
 CC asthma and leukemias, and agonists can be used in the treatment
 CC of e.g. anaemias, neutropenias, thrombocytopenia, cancer, obesity
 CC viral diseases, growth retardation or other diseases characterised
 CC by insufficient Stat activity.
 XX
 XX Sequence 750 AA:

SQ Sequence 750 AA;

Query Match	50.18	Score 2017	DB 19	Length 750
Best Local Similarity	52.58	Pred. No. 1.2e-165		
Matches 389	Conservative 145	Mismatches 193	Indels 14	Gaps 9

```

0Y 1 MAONMLODOPRTRYKOLJOLHOLYSDPEFMEHROGLAPWIESDOMAAASKBESHATVFNHNL 60
0Y 1 msqweylejdaqlflegvnyqlyddsfmeitfrylaqwltekdqwehandaufatitfhnd 60
Db 1 msqweylejdaqlflegvnyqlyddsfmeitfrylaqwltekdqwehandaufatitfhnd 60
0Y 61 LGEIDOOYRSRELOESNVLVYUHNLRRIKOFLOSRYLEKEMELARIYARLCEESRLDTPAA 120
Db 61 lsglddgyrsfelenfllghnrtsknldgnrfedoprigmsllyscelkeerklendq 120
0Y 121 TAAOGGOGANHPPTAAVTEKOOOLEQHLODYRRVODLEOKMKVVENLODDFDFNYTKL 180
Db 121 rfnq--agsnqsgtswmdkqekldskvrvnwkdkwmcshelkshdiedydefekctf- 177
0Y 181 SOGDODLNGNNOSTROCKMOOLEMLTALDOMRKSIVSELAGLISAMHYOKTLTHBEL 240
Db 178 -qnehecnvayksdqbkqegllkmylm|dnkrtevhklll|ellnyleltna|lndel 236
0Y 241 ADMKRRPEIACIGSPNLCIDRLBNMWTSLAESOLQOTQOIKLEELDOOKYKGPDIYO 300
Db 237 vewkrtsqsgagppnacd|qmwftvices|qgvrvq|kklleelqklytyehnoplt 296
0Y 301 HRPMEERIVELFRNLKMSAVVEROPCMPHRPHVIAKTGOVPTTKRVILKYKPREIN 360
Db 297 nkqytwdfctfsl|eqql|sgstfvereqpmpubhrpqrp|vltktygfvtk|tr|vklqelny 356
0Y 361 OLKTIYVCIDKOSGOVAAALRGSRKFNLLICTNNKYNMMESSNNGSISAPFKHLLTDEQSCN 420
Db 357 nkkrvvtldkdvnenntvckgfrfkn|lgythkmmesttrngs|aaetnrh|q|keq--n 414
0Y 421 GGRANCDASLIVTEELHLITFEVEYHOGKLIDETSHLPVVVISN|COMPNMAWASILMY 480
Db 415 agtrtnegp|lvteel|hs|fsteq|cpgy|vdiets|lpvvvisnvsq|psgwast|wy 474
0Y 481 NMLNPNKNNVFPFKP|IGTMDXDAEVI|SWQSS|TTRKGIS|EOLU|TLAERKL|GPGVYNS 540
Db 475 nmlvaepnrnlsf|llppcarwaq|sevw|swgtsfsv|ktrg|lrvndq|lmg|kell|gpnaspd 534
0Y 541 GCQUTMAKFCCKENNAKGFSFWWLDN|IDLAVKYU|LALNNEG|YIMG|ISKERERAL|ST 600
Db 535 gl-lpwtrfcikenndnfrpwlies|lel|khh|lpw|ndcgng|f|sketeral|xd 593
0Y 601 KPPEOTFLIARESSSEKEGCVTF|TWEEKDIS--GKTJOLQ|SEVPTKQOLNNMS|FAT|INGYKI 659
Db 594 qgqpf|fl|trfsees|reda|t|l|twers|qndqep|f|havery|lkekel|sav|fod|l|nrvy 653

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 19, 2002, 15:21:58 ; Search time 56.45 Seconds
(without alignments)
1995.212 Million cell updates/sec

Title: US-08-212-185-12
Perfect score: 4029
Sequence: 1 MAQNNQLOQLDTRFKQLKHLQ.....QFESLTFMDLTSQCATSPM 770

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_17:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4006	99.4	770	4	014916
2	3891.5	99.1	769	4	09BW54
3	3839.5	95.3	769	13	09PVX8
4	3474	86.2	806	13	093599
5	3373.5	83.7	767	13	013133
6	2038	50.6	749	11	09QXK0
7	2026	50.3	749	11	09QXJ2
8	2006	49.8	712	11	09QXK4
9	1901	47.2	754	13	013131
10	1891	46.9	754	13	013132
11	1786.5	44.3	749	13	093598
12	1597.5	39.7	394	13	09QXJ8
13	1287	31.9	922	11	09QXJ2
14	1284.5	31.9	925	11	09QXJ4
15	873	21.7	784	13	09QXJ4
16	867	21.5	787	13	09QXJ4
17	865.5	21.5	786	11	09QXJ1
18	834	20.7	786	11	09QXJ1
19	663.5	16.5	592	13	09PWP7

20	566	14.0	679	4	09BOD2	09bq2 homo sapien
21	532	13.2	722	5	097164	097164 anopheles g
22	473.5	11.8	156	4	09UDL5	09ud5 homo sapien
23	423	10.5	703	5	09NAD6	09nad6 caenorhabdl
24	389.5	9.7	195	6	09GKY9	09gky9 sus scrofa
25	342	8.5	155	4	09UDL4	09ud4 homo sapien
26	299.5	7.4	144	11	070428	070428 rattus norv
27	297	7.4	111	11	070406	070406 rattus norv
28	292	7.2	56	6	09N0J4	09n0j4 bos laurus
29	280.5	7.0	177	6	09N0E5	09n0e5 bos laurus
30	273	6.8	51	11	09NML3	09nml3 mus musculu
31	266	6.6	51	4	09BXH2	09bxh2 homo sapien
32	256.5	6.4	85	6	029340	029340 sus scrofa
33	251.5	6.2	141	11	070429	070429 rattus norv
34	248.5	6.2	230	11	09R0X9	09r0x9 mus musculu
35	224	5.6	43	11	09QVR4	09qvr4 rattus sp.
36	223.5	5.5	205	11	09R0X8	09r0x8 mus musculu
37	215	5.3	85	6	029356	029356 sus scrofa
38	211.5	5.2	929	5	09BLX2	09blx2 dictyostell
39	178	4.4	43	11	09QVR3	09qvr3 rattus sp.
40	177.5	4.4	707	5	000910	000910 dictyostell
41	164.5	4.1	1356	4	014707	014707 homo sapien
42	161	4.0	2041	4	09HAT7	09hat7 homo sapien
43	161	4.0	2047	4	09UH61	09uh61 homo sapien
44	159.5	4.0	2101	4	014981	014981 homo sapien
45	156.5	3.9	1300	4	013999	013999 homo sapien

ALIGNMENTS

RESULT	1	INITIALINARY;	PRT;	770 AA.
ID	014916	014916		
AC	014916			
DT	01-JAN-1998 (TREMBL)	05, Created		
DT	01-AUG-1999 (TREMBL)	11, Last sequence update		
DT	01-JUN-2001 (TREMBL)	17, Last annotation update		
DE	TRANSCRIPTION FACTOR (SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION			
DE	3).			
GN	STAT3			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	11			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=8826260; PubMed=9630560;			
RA	Della Pietra L., Bressan A., Pezzotti A., Serlupi-Crescenzi O.,			
RT	*Highly conserved amino-acid sequence between murine STAT3 and a			
RT	revised human STAT3 sequence.*;			
RL	Gene 213:119-124(1998).			
RN	12			
RP	SEQUENCE OF 564-704 FROM N.A.			
RC	TISSUE=HEPATOMA;			
RA	Della Pietra L., Bressan A., Pezzotti A.R., Serlupi-Crescenzi O.,			
RL	Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AJ012463; CAA10032.1;			
DR	EMBL: AF029311; ABB84254.1;			
DR	HSSP: P42227; IBGL			
DR	InterPro: IPR000980; SH2			
DR	InterPro: IPR001217; STAT			
DR	Pfam: PF00017; SH2; 1.			
DR	Pfam: PF01017; STAT; 1.			
DR	SMART: SM00252; SH2; 1.			
DR	PROSITE: PSS0001; SH2; 2.			
SO	SEQUENCE 770 AA; 88067 MW; 6C00632211C8012D CRC64;			

Query Match 99.4%; Score 4006; DB 4; Length 770;
Best Local Similarity 99.4%; Pred. No. 3.2e-263;
Matches 765; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 19, 2002, 15:25:48 : Search time 16.05 Seconds
(without alignments)
3654.480 Million cell updates/sec

Title: US-08-212-185-12

Sequence: 770
1 MAQWNOLOQUDTRYLKQLHQ.....QFESLTEDMDLTSECATSPM 770

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : PIR68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	442	57.4	770	2	I49508	ISGF3 p91-related
2	221	28.7	770	2	A54444	DNA-binding protei
3	12	1.6	739	2	A46159	Interferon-depende
4	11	1.4	57	2	S63680	signal transducer
5	11	1.4	65	2	S63679	signal transducer
6	11	1.4	748	2	A56047	gamma-interferon a
7	11	1.4	851	2	A46160	interferon alpha-1
8	9	1.2	44	2	T26893	hypothetical prote
9	9	1.2	72	2	T26170	hypothetical prote
10	8	1.0	293	2	G82180	transcription regu
11	8	1.0	663	2	S67259	transcription regu
12	8	1.0	728	2	F84279	MEI protein - yea
13	8	1.0	786	2	I49274	amino acid transp
14	8	1.0	793	2	S54772	mammary gland fac
15	8	1.0	794	2	G02317	mammary gland fac
16	8	1.0	794	2	S55527	transcription acti
17	8	1.0	837	2	I57557	mammary g.-and fac
18	8	1.0	837	2	I57557	DNA-binding protei
19	8	1.0	848	1	A54740	interleukin-4-indu
20	8	1.0	1088	1	P1XRBR	inner layer protei
21	8	1.0	1088	1	P1XRBR	inner layer protei
22	8	1.0	1088	1	P1XRBR	inner layer protei
23	8	1.0	1088	2	S13558	VP1 protein - porc
24	7	0.9	49	2	S63682	signal transducer
25	7	0.9	64	2	S63681	signal transducer
26	7	0.9	77	2	H70642	probable ribosomal
27	7	0.9	94	2	F31844	spda protein - Str
28	7	0.9	111	2	T17132	hypothetical prote
29	7	0.9	128	2	H70435	hypothetical prote

ALIGNMENTS

RESULT	1	ALIGNMENTS
I49508	ISGF3 p91-related transcription factor - mouse	
C:Species: Mus musculus (house mouse)		
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 01-Dec-2000		
C:Accession: I49508; I49009		
R:Akira, S.; Nishio, Y.; Inoue, M.; Wang, X.J.; Wei, S.; Matsusaka, T.; Yoshida, K.; Cell 77, 63-71, 1994		
A:Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related		
A:Reference number: A54444, PMID:94208062		
A:Accession: I49508		
A>Status: preliminary; translated from GB/EMBL/DBJ		
A:Molecule type: mRNA		
A:Residues: 1-770 <RES>		
A:Cross-references: GB:I29278; NID:q476715; PIDN:AAA37254.1; PID:q476716		
R:Raz, R.; Durbin, J.E.; Levy, D.E.		
J. Biol. Chem. 269, 24391-24395, 1994		
A:Title: Acute phase response factor and additional members of the interferon-stimula		
A:Reference number: I49009; PMID:95014185		
A:Accession: I49009		
A>Status: preliminary; translated from GB/EMBL/DBJ		
A:Molecule type: mRNA		
A:Residues: 1-393, 'M', 395-700, 702-770 <RES>		
A:Cross-references: EMBL:U08378; NID:q473889; PIDN:AAA56668.1; PID:q473890		
C:Genetics: APRF		
C:Superfamily: human signal transducer and transcription activator STAT5A		
Query Match	57.4%; Score 442; DB 2; Length 770;	
Best Local Similarity	99.6%; Pred. No. 0;	
Matches 742; Conservative 0; Mismatches 3; Indels 0; Gaps 0;		
OY	26 FPEMLRQFLAPWIESQDMAYAAKESHATLVFNHLLGEIDQYSRFLOESNVLYQNLRR 85	Alu RNA-binding pr
		signal recognition
Db	26 FPEMLRQFLAPWIESQDMAYAAKESHATLVFNHLLGEIDQYSRFLOESNVLYQNLRR 85	cell division cont
		hypothetical prote
OY	86 IKQFLOSRYLEKPEMEIARIYARCLMESRLLOTAATAAOGGGAHPPTAAVTEKQOMLE 145	hypothetical prote
		hypothetical prote
Db	86 IKQFLOSRYLEKPEMEIARIYARCLMESRLLOTAATAAOGGGAHPPTAAVTEKQOMLE 145	hypothetical prote
		hypothetical prote
OY	146 QHLDQVKKRVQDLEQKKVVENLDQDDDFYKTLKSGDMQDINGNNOSTYRQMOOLEQ 205	mitotic feedback c
		hypothetical prote
Db	146 QHLDQVKKRVQDLEQKKVVENLDQDDDFYKTLKSGDMQDINGNNOSTYRQMOOLEQ 205	hypothetical prote
		hypothetical prote
OY	206 MLTALDDMRRSIVSELAGLISAMEYVOKTLTDELDADWKRREPITACTGSPNITCLDRLN 265	C-8 sterol isomera
		hypothetical prote
Db	206 MLTALDDMRRSIVSELAGLISAMEYVOKTLTDELDADWKRREPITACTGSPNITCLDRLN 265	hypothetical prote
		hypothetical prote
OY	266 WITSLSAESQLOTROQIKKLELQOKVSKGDPVIVQHPMLEEITVELFRLMKSAVYVER 325	hypothetical prote
		hypothetical prote
Db	266 WITSLSAESQLOTROQIKKLELQOKVSKGDPVIVQHPMLEEITVELFRLMKSAVYVER 325	accessory gene reg

OY	326	OPCMHBPDRPLVKTGQFETTKRLLYKRPPELNYOLKIKVCIIDKXSGVAAALRGRKPN	385
Db	326	OPCMHBPDRPLVKTGQFETTKRLLYKRPPELNYOLKIKVCIIDKXSGVAAALRGRKPN	385
OY	386	ILGTNTKIVMNNSEENSGSLSAEFKHLTLRQRCGNGRANCADSLVTEELHITFEYEV	445
Db	386	ILGTNTKIVMNNSEENSGSLSAEFKHLTLRQRCGNGRANCADSLVTEELHITFEYEV	445
OY	446	YHOGSLKILLEHSLRPVYVYISNICMPNMAASILMYNMLTNPNVNFPTKPIGTMDOYA	505
Db	446	YHOGSLKILLEHSLRPVYVYISNICMPNMAASILMYNMLTNPNVNFPTKPIGTMDOYA	505
OY	506	EVLMSQFSSTTKRGSLIEQLTLTAEKILLGCVNGVSGOITMAFKCEKNNAKGFSPWVL	565
Db	506	EVLMSQFSSTTKRGSLIEQLTLTAEKILLGCVNGVSGOITMAFKCEKNNAKGFSPWVL	565
OY	566	DNIIDLVRKYLLALMNBGYIMGFISKREKRAILSTKPKGTELLRFSESSKEGVFTTWE	625
Db	566	DNIIDLVRKYLLALMNBGYIMGFISKREKRAILSTKPKGTELLRFSESSKEGVFTTWE	625
OY	626	KDIOISKTQIOQVEPEPTKQOLNNMSFAELIMGYKIMDATNITVSPVLYLIPDIPKEAFCK	685
Db	626	KDIOISKTQIOQVEPEPTKQOLNNMSFAELIMGYKIMDATNITVSPVLYLIPDIPKEAFCK	685
OY	686	YCRPESOEHPADPGSAAPYLKTKFICVPTPTCSNTIDLPMSPRTDLSLMOFGNNGEAE	745
Db	686	YCRPESOEHPADPGSAAPYLKTKFICVPTPTCSNTIDLPMSPRTDLSLMOFGNNGEAE	745
OY	746	PSAGQGFESLTFDDMLTISECATSPM	770
Db	746	PSAGQGFESLTFDDMLTISECATSPM	770

RESULT 2
A54444
DNA-binding protein APRF - human
C:Species: homo sapiens (man)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 28-Jul-2000
C:Accession: A54444
R:Akita, S.; Nishio, Y.; Inoue, M.; Wang, X.-J.; Wei, S.; Matsusaka, T.; Yoshida, K.; Sud-
Cell 77, 63-71, 1994
A:Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related tra-
A:Reference number: A54444; MUID:94208062
A:Accession: A54444
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-770 <RES>
A:Cross-references: GB:L29277; NID:g475788; PID:g475789
C:Genetics:
A:Gene: GDB:STAT3; APRF
A:Cross-references: GDB:356950
A:Map position: 17q21-17q21
C:Superfamily: human signal transducer and transcription activator STAT5A
C:Keywords: DNA binding; transcription factor

Query Match	28.7%	Score 221	DB 2	Length 770
Best Local Similarity	100.0%	Pred. NC	1.1e-220	
Matches 221	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	26	FPELJROGLAMIESODPAAVAAKSESHATLVFNHLJGEIDOOYSRFLJESNVLVYOHNLJR	85
Db	26	FPELJROGLAMIESODPAAVAAKSESHATLVFNHLJGEIDOOYSRFLJESNVLVYOHNLJR	85
QY	86	IKOFJOSYXLEKKPHEIATVACRAMESSRLLOTPATAOOGGCAHPTAAYVTEKOOOLE	145
Db	86	IKOFJOSYXLEKKPHEIATVACRAMESSRLLOTPATAOOGGCAHPTAAYVTEKOOOLE	145
QY	146	QHLDOVRRVODLEOKKAVENLODDFDENYKTLKLSOGDMODLNGNOSVYTRKMOOLE	205
Db	146	QHLDOVRRVODLEOKKAVENLODDFDENYKTLKLSOGDMODLNGNOSVYTRKMOOLE	205

Oy	206	MLTALDQMRRSIVSELGLLSAMEYVOKTLTDEELADWKR	R	246
Dd	206	MLTALDQMRRSIVSELGLLSAMEYVOKTLTDEELADWKR	R	246

RESULT 3
AA6150

Interferon-dependent positive-acting transcription factor ISGF-3 91k chain - human
 C:Species: Homo sapiens (man)
 C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 28-Jul-2000
 C:Accession: A46159
 R:Schindler, C.; Fu, X.-Y.; Impirola, T.; Aebersold, R.; Darnell Jr., J.E.
 Proc. Natl. Acad. Sci. U.S.A. 89, 7836-7839, 1992
 A:Title: Proteins of transcription factor ISGF-3: one gene encodes the 91- and 84-kDa
 A:Reference number: A46159; MUID:92166557
 A:Accession: A46159
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: nucleic acid; protein
 A:Residues: 1-739 <SCCH>
 A:Experimental source: HeLa cells
 A:Note: sequence extracted from NCBI backbone (NCBIP:110818)
 C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match	1.6%	Score 12;	DB 2;	Length 739;
Best Local Similarity	100.0%;	Pred. No. 0.0013;		
Matches 12; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

```
QY      585  IMGFISKERERA 596
         |||||
Db      567  IMGFISKERERA 578
```

RESULT 4
S63680
signal transducer and activator of transcription 2 type c - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #texL_change 02-Sep-2000
C:Accession: S63680
R:Sugiyama, T.; Nishio, Y.; Kishimoto, T.; Akira, S.
FEBS Lett. 381, 191-194, 1996
A>Title: Identification of alternative splicing form of Stat2.
A:Reference number: S63679; MUID:96176320
A:Accession: S63680
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: mRNA
A:Residues: 1-57 <SUG>
A:Cross-references: GB:S81491; MID:q1478435; PIDN:AAB36226.1; PID:q1478435;
C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match	1.4%	Score 11	DB 2	Length 57
Best Local Similarity	100.0%	Pred. No.	0.0015	
Matches 11	Conservative 0	Mismatches 0	Indels 0	Gaps 0

```
QY      604 GTFLRFSESS 614
        |||||
Db      1 GTFLRFSESS 11
```

RESULT 5
S63679
signal transducer and activator of transcription 2 type a - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 02-Sep-2000
C:Accession: S63679
R:Sugiyama, T.; Nishio, Y.; Kishimoto, T.; Akira, S.
FEBS Lett. 381, 191-194, 1996
A:Title: Identification of alternative splicing form of Stat2.
A:Reference number: S63679; MUID:96176320
A:Accession: S63679
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: mRNA

A:Residues: 1-65 <SUG>
A:Cross-references: GB:S81491; NID:q1478435; PIDN:AA636225.1; PID:q1478436
C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 1.4%; Score 11; DB 2; Length 65;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 604 GTFLRFSS 614
|||||
DB 1 GTFLRFSS 11

RESULT 6
A56047
gamma-interferon activation site-binding protein Stat4 - mouse

C:Species: Mus musculus (house mouse)
C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 28-Jul-2000
C:Accession: A56047
R:Yamamoto, K.; Queille, F.W.; Thierfelder, W.E.; Kreider, B.L.; Gilbert, D.J.; Jenkins, M.J. Cell. Biol. 14, 4342-4349, 1994
A:Title: Stat4, a novel gamma interferon activation site-binding protein expressed in e
A:Reference number: A56047; MUID:94277038
A:Accession: A56047
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-748 <YAM>
A:Cross-references: GB:U09351; NID:q509502; PIDN:AAA19692.1; PID:q509503
C:Superfamily: human signal transducer and transcription activator STAT5A
C:Keywords: DNA binding; phosphoprotein

Query Match 1.4%; Score 11; DB 2; Length 748;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 603 PGFLRFSS 613
|||||
DB 592 PGFLRFSS 602

RESULT 7
A46160
interferon alpha-induced transcription activator ISGF-3, 113K chain - human

N:Alternate names: stat2 protein
C:Species: Homo sapiens (man)
C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 18-Aug-2000
C:Accession: A46160; S71908; S53873
R:Fu, X.Y.; Schindler, C.; Impirota, T.; Aebersold, R.; Darnell Jr., J.E.
Proc. Natl. Acad. Sci. U.S.A. 89, 7840-7843, 1992
A:Title: The proteins of ISGF-3, the interferon alpha-induced transcriptional activator,
A:Reference number: A46160; MUID:92366558
A:Accession: A46160
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA; protein
A:Residues: 1-851 <FU>
A>Note: sequence extracted from NCBI backbone (NCBIP:110820)
R:Yan, R.; Qureshi, S.; Zhong, Z.; Wen, Z.; Darnell, J.E.
submitted to the EMBL Data Library, December 1994
A:Reference number: S71908
A:Accession: S71908
A:Molecule type: DNA
A:Residues: 1-851 <YAN>
A:Cross-references: EMBL:U18671; NID:q1293919; PIDN:AA98760.1; PID:q1293920
R:Yan, R.; Qureshi, S.; Zhong, Z.; Wen, Z.; Darnell Jr., J.E.
Nucleic Acids Res. 23, 459-463, 1995
A:Title: The genomic structure of the STAT genes: multiple exons in coincident sites in
A:Reference number: S53873; MUID:95192056
A:Accession: S53873
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-196;392-591;684-730 <YAM>

A:Cross-references: EMBL:U18671
C:Genetics:
A:Gene: stat2
A:Introns: 44/2; 95/3; 127/3; 157/3; 183/1; 211/3; 261/2; 314/2; 345/2; 372/2;
C:Superfamily: human signal transducer and transcription activator STAT5A
C:Keywords: signal transduction; transcription regulation

Query Match 1.4%; Score 11; DB 2; Length 851;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 604 GTFLRFSS 614
|||||
DB 596 GTFLRFSS 606

RESULT 8
T26893
hypothetical protein Y44A6D.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26893
R:Ainscough, R.
submitted to the EMBL Data Library, June 1998
A:Reference number: Z20282
A:Accession: T26893
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-44 <WIL>
A:Cross-references: EMBL:AL023842; PIDN:CAA19514.1; GSPDB:GN00023; CESP:Y44A6D.1
A:Experimental source: clone Y44A6D
C:Genetics:
A:Gene: CESP:Y44A6D.1
A:Map position:
A:Introns: 30/1

Query Match 1.2%; Score 9; DB 2; Length 44;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 282 KKLDELQOK 290
|||||
DB 33 KKLDELQOK 41

RESULT 9
T26170
hypothetical protein W04G5.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26170
R:Kershaw, J.
submitted to the EMBL Data Library, March 1997
A:Reference number: Z20164
A:Accession: T26170
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-72 <WIL>
A:Cross-references: EMBL:Z93391; PIDN:CAB07680.1; GSPDB:GN00019; CESP:W04G5.3
A:Experimental source: clone W04G5
C:Genetics:
A:Gene: CESP:W04G5.3
A:Map position: 1
A:Introns: 20/1

Query Match 1.2%; Score 9; DB 2; Length 72;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 282 KKLDELQOK 290

Db 23 KKEELLOOK 31
RESULT 10
G82180
transcription regulator lysr family VC1588 [imported] - *Vibrio cholerae* (strain N16961)
C:Species: *Vibrio cholerae*
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: G82180
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
Charson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.
I. R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: A82035; MUID:2040683
A:Accession: G82180
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-293 <HEI>
A:Cross-references: GB:AE004236; GB:AE003852; NID:g9656095; PIDN:AAF94742.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1588
A:Map position: 1
C:Superfamily: Pseudomonas putida regulatory protein catr

Query Match 1.0%; Score 8; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 279 OQIKLEE 286
Db 35 OQIKLEE 42
RESULT 11
S67259
MNE1 protein - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: protein 06353; protein YOR350C
C:Species: *Saccharomyces cerevisiae*
C>Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 21-Jul-2000
C:Accession: S67259; S67262; S20175; S67404; S19073
R:Goffeau, A.; Purnelle, B.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67246
A:Accession: S67259
A:Molecule type: DNA
A:Residues: 1-663 <GOR>
A:Cross-references: EMBL:275258; NID:g1420761; PID:g1420762; MIPS:YOR350C
A:Experimental source: strain S288C
R:Deilus, H.; Hebling, U.; Hofmann, B.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67261
A:Accession: S67262
A:Molecule type: DNA
A:Residues: 1-663 <DEI>
A:Cross-references: EMBL:275258; NID:g1420761; PID:g1420762; MIPS:YOR350C
A:Experimental source: strain S288C
R:Leem, S.H.; Ogawa, H.
Nucleic Acids Res. 20, 449-457, 1992
A:Title: The MRP4 gene encodes a novel protein kinase homologue required for meiotic recombination
A:Reference number: S20174; MUID:92158649
A:Accession: S20175
A:Molecule type: DNA
A:Residues: 1-219 <LEE>
A:Cross-references: EMBL:X63112; NID:g3968; PIDN:CAA44826.1; PID:g3970
R:Purnelle, B.; Goffeau, A.
submitted to the EMBL Data Library, February 1996
A:Description: Nucleotide sequence analysis of a 40 kb segment on the right arm of yeast
ome 1 genes.
A:Reference number: S67392

A:Accession: S67404
A:Molecule type: DNA
A:Residues: 1-663 <PUR>
A:Cross-references: EMBL:X95720; NID:g1199839; PID:g1199852
C:Genetics:
A:Gene: SCD:MNE1
A:Cross-references: SCD:S0005877; MIPS:YOR350C
A:Map position: 15R
C:Superfamily: *Saccharomyces cerevisiae* MNE1 protein

Query Match 1.0%; Score 8; DB 2; Length 663;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 590 SKERERAI 597
Db 249 SKERERAI 256
RESULT 12
F84279
amino acid transporter [imported] - *Halobacterium* sp. NRC-1
C:Species: *Halobacterium* sp. NRC-1
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: F84279
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laaky
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Ja
Jung, K.H.; Alam, M.; Freltas, F.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Edhardt, H.; Lowe, T.M.;
A:Title: Genome sequence of *Halobacterium* species NRC-1.
A:Reference number: A84160; MUID:20504483
A:Accession: F84279
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-728 <STO>
A:Cross-references: GB:AE004437; NID:g10580770; PIDN:AAG19602.1; GSPDB:GN00138
C:Genetics:
A:Gene: yhdg

Query Match 1.0%; Score 8; DB 2; Length 728;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 220 ELAGLLSA 227
Db 653 ELAGLLSA 660
RESULT 13
I49274
mammary gland factor - mouse
N:Alternate names: STAT5 protein homolog p80
C:Species: *Mus musculus* (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 01-Dec-2000
C:Accession: I49274; S54773; S54727
R:Jin, X.; Robinson, G.W.; Gouilleux, F.; Groner, B.; Hennighausen, L.
Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995
A:Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involve
A:Reference number: I49273; MUID:96004632
A:Accession: I49274
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-786 <RES>
A:Cross-references: EMBL:U21110; NID:g747973; PIDN:AAC52282.1; PID:g747974
R:Jin, A.L.F.; Makao, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.
EMBO J. 14, 1166-1175, 1995
A:Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin
A:Reference number: S54772; MUID:95237198
A:Accession: S54773
A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA
A:Residues: 1-432, 'E', 434-786 <MU1>
A:Cross-references: EMBL:248539; NID:g758635; PIDN:CAA8420.1; PID:g758636
R:Azam, M.; Erdjument-Bromage, H.; Kreider, B.L.; Xie, M.; Queller, F.; Basu, R.; Saris, E.M.O. J. 14, 1407-1411, 1995
A:Title: Interleukin-3 signals through multiple isoforms of Stat5.
A:Reference number: S54725; MUID:95246733
A:Accession: S54727
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-432, 'E', 434-786 <AZA>
C:Genetics:
A:Gene: Stat5b
C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 1.0%; Score 8; DB 2; Length 786;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 604 GTFLLRFS 611
|||||
DB 613 GTFLLRFS 620

RESULT 14
S54772
mammary gland factor - mouse
N:Alternate names: stat5 protein
C:Species: Mus musculus (house mouse)
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 28-Jul-2000
C:Accession: S54772; 149273
R:Mu, A.L.F.; Wako, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.
EMBO J. 14, 1166-1175, 1995
A:Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin
A:Reference number: S54772; MUID:95237198
A:Accession: S54772
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-793 <MU1>
A:Cross-references: EMBL:248538; NID:g758633; PIDN:CAA8419.1; PID:g758634
R:Liu, X.; Robinson, G.W.; Guilleux, F.; Groner, B.; Hennighausen, L.
Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995
A:Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involved in
A:Reference number: 149273; MUID:96004632
A:Accession: 149273
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-793 <RES>
A:Cross-references: EMBL:U21103; NID:g747971; PIDN:AAA80590.1; PID:g747972
C:Genetics:
A:Gene: Stat5a
C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 1.0%; Score 8; DB 2; Length 793;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 604 GTFLLRFS 611
|||||
DB 613 GTFLLRFS 620

RESULT 15
G02317
transcription activator stat5A - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 28-Jul-2000
C:Accession: G02317
R:Liu, J.
submitted to the EMBL Data Library, December 1995
A:Reference number: H01043

A:Accession: G02, 7
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-794 <LIN>
A:Cross-references: EMBL:043385; NID:g1151169; PIDN:AA806589.1; PID:g1151170
C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 1.0%; Score 8; DB 2; Length 794;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 604 GTFLLRFS 611
|||||
DB 613 GTFLLRFS 620

Search completed: March 19, 2002, 15:29:01
Job time: 193 s.

RP SEQUENCE FROM N.A.
RA Nishinakamura R., Matsumoto Y., Matsuda T., Arizumi T., Heike T.,
KA Asashima M., Yokota T.,
RT "Activation of Stat3 by cytokine receptor gp130 ventralizes Xenopus
embryos independently of BMP-4.";
KL Dev. Biol. 0:0-0(1999).
DR EMBL: AB017701; BAA86061.1; -;
DR HSSP: P42227; 1BG1.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SQ SEQUENCE 769 AA: 87974 MW: 0905C03263303069 CRC64;

Query Match 95.3%; Score 3839.5; DB 13; Length 769;
Best local similarity 95.1%; Pred. No. 6e-252;
Matches 733; Conservative 19; Mismatches 16; Indels 3; Gaps 2;

QY 1 MAQNNLOQLDRTYRLKQHLQYSDTFPMELRQFLAPWIESODMAVYASKESHATLVFHNL 60
DB 1 MAQNNLOQLDRTYRLKQHLQYSDTFPMELRQFLAPWIESODMAVYASKESHATLVFHNL 60
QY 61 LGEIDQYSRFLQESNVLYOHNLRRIKQFLQSRYLEKPMETARIVARCLWESRLQTAA 120
DB 61 LGEIDQYSRFLQESNVLYOHNLRRIKQFLQSRYLEKPMETARIVARCLWESRLQTAA 120
QY 121 TAAQGGGANHPPTAAVYTEKQMLEOHLDYRKRYODLEQKMKVENVLQDDFDFNYKTLK 180
DB 121 TAAQGGGANHPPTAAVYTEKQMLEOHLDYRKRYODLEQKMKVENVLQDDFDFNYKTLK 180
QY 121 AAAQGGPASHIPNAVYTEKQMLEOHLDYRKRYODLEQKMKVENVLQDDFDFNYKTLK 180
DB 121 AAAQGGPASHIPNAVYTEKQMLEOHLDYRKRYODLEQKMKVENVLQDDFDFNYKTLK 180
QY 181 SOGMDOLNCGNSVTRQKMOOLEQMLTALDOMRSIVSELAGLSAMEYVQKLTDEEL 240
DB 181 SOGMDOLNCGNSVTRQKMOOLEQMLTALDOMRSIVSELAGLSAMEYVQKLTDEEL 240
QY 181 SOSLSLSELNGNSVTRQKMOOLEQMLTALDOMRSIVSELAGLSAMEYVQKLTDEEL 240
DB 181 SOSLSLSELNGNSVTRQKMOOLEQMLTALDOMRSIVSELAGLSAMEYVQKLTDEEL 240
QY 241 ADMKRRPELIACIGPPNICDLRENMWITSLAESQLTRQOIKKLEELQOKVSYKGDPIVQ 300
DB 241 ADMKRRPELIACIGPPNICDLRENMWITSLAESQLTRQOIKKLEELQOKVSYKGDPIVQ 300
QY 301 HRPLLEERIVELFRLNLSAFVVEROPCMHPRPLVITKGVQFTTKVRLLVKFPPELNY 360
DB 301 HRPLLEERIVELFRLNLSAFVVEROPCMHPRPLVITKGVQFTTKVRLLVKFPPELNY 360
QY 301 HRPLLEERIVELFRLNLSAFVVEROPCMHPRPLVITKGVQFTTKVRLLVKFPPELNY 360
DB 301 HRPLLEERIVELFRLNLSAFVVEROPCMHPRPLVITKGVQFTTKVRLLVKFPPELNY 360
QY 361 QLKIKVCIDKDSGDVAALRGSRRKFNILGTNTKVMNEESNNGSLSAEFKHLTLREQRCGN 420
DB 361 QLKIKVCIDKDSGDVAALRGSRRKFNILGTNTKVMNEESNNGSLSAEFKHLTLREQRCGN 420
QY 421 GGRANCDASLVTEBELHLTFETEVYHOGKIDLETHSLRPVVVISNICOMPAASILMY 480
DB 421 GGRANCDASLVTEBELHLTFETEVYHOGKIDLETHSLRPVVVISNICOMPAASILMY 480
QY 481 NMLNLRKNVNFETKPPIGTMDQVAEVLWQFSSTTRKGLSIEQLTLAEKLLRPGVNS 540
DB 481 NMLNLRKNVNFETKPPIGTMDQVAEVLWQFSSTTRKGLSIEQLTLAEKLLRPGVNS 540
QY 541 GCOITWAKFCKENNAAGGFSFWWLNDIIDLKYYILALMNGYIMGFISKEERAILST 600
DB 541 GCOITWAKFCKENNAAGGFSFWWLNDIIDLKYYILALMNGYIMGFISKEERAILST 600
QY 601 KPPGTFILRSESSKEGVTFTWEKDISGKTQISVEPYTKQOLNNMSPFAEILMGYKIM 660
DB 601 KPPGTFILRSESSKEGVTFTWEKDISGKTQISVEPYTKQOLNNMSPFAEILMGYKIM 660
QY 661 DATNIVLSPVLYLPDIKREBAFGKICRPESQHPH-ADPGSAAPYLKTKFICVPTPTCS 719
DB 661 DATNIVLSPVLYLPDIKREBAFGKICRPESQHPH-ADPGSAAPYLKTKFICVPTPTCS 719
QY 720 NTIDLPSPRTIDSLMOFGNNGCAEPGAGPESLTFDMDLTSECATSPM 770
DB 720 NTIDLPSPRTIDSLMOFGNNGCAEPGAGPESLTFDMDLTSECATSPM 770
QY 721 STIDLPSPRTIDSLMOFGNNGCAEPGAGPESLTFDMDLTSECATSPM 769
DB 721 STIDLPSPRTIDSLMOFGNNGCAEPGAGPESLTFDMDLTSECATSPM 769

RESULT 4
ID 093599 PRELIMINARY: PRT: 806 AA.
AC 093599;
DT 01-NOV-1998 (TREMBLrel. 08, created)
DT 01-NOV-1998 (TREMBLrel. 08, last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE TRANSCRIPTION FACTOR.
GN STAT3.
OS Brachydanio rerio (Zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Rasbora; Danio.
OX NCBI_TaxId=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Oates A.C.;
RL Thesis (1998), University of Melbourne, Australia.
DR EMBL: AJ005693; CAA06677.1; -;
DR HSSP: P42227; 1BG1.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SQ SEQUENCE 806 AA: 92151 MW: 74RC4EA401C3C942 CRC64;

Query Match 86.2%; Score 3474; DB 13; Length 806;
Best local similarity 85.1%; Pred. No. 3.8e-227;
Matches 678; Conservative 43; Mismatches 40; Indels 36; Gaps 11;

QY 1 MAQNNLOQLDRTYRLKQHLQYSDTFPMELRQFLAPWIESODMAVYASKESHATLVFHNL 60
DB 1 MAQNNLOQLDRTYRLKQHLQYSDTFPMELRQFLAPWIESODMAVYASKESHATLVFHNL 60
QY 61 LGEIDQYSRFLQESNVLYOHNLRRIKQFLQSRYLEKPMETARIVARCLWESRLQTAA 120
DB 61 LGEIDQYSRFLQESNVLYOHNLRRIKQFLQSRYLEKPMETARIVARCLWESRLQTAA 120
QY 61 LGEIDQYSRFLQESNVLYOHNLRRIKQFLQSRYLEKPMETARIVARCLWESRLQTAA 120
DB 61 LGEIDQYSRFLQESNVLYOHNLRRIKQFLQSRYLEKPMETARIVARCLWESRLQTAA 120
QY 121 TAAQGGGANHPPTAAVYTEKQMLEOHLDYRKRYODLEQKMKVENVLQDDFDFNYKTLK 180
DB 121 TAAQGGGANHPPTAAVYTEKQMLEOHLDYRKRYODLEQKMKVENVLQDDFDFNYKTLK 180
QY 120 TTAQDDQVAAHPPTVYTEKQMLEOHLDYRKRYODLEQKMKVENVLQDDFDFNYKTLK 179
DB 120 TTAQDDQVAAHPPTVYTEKQMLEOHLDYRKRYODLEQKMKVENVLQDDFDFNYKTLK 179
QY 181 SOGMDOLNCGNSVTRQKMOOLEQMLTALDOMRSIVSELAGLSAMEYVQKLTDEEL 240
DB 181 SOGMDOLNCGNSVTRQKMOOLEQMLTALDOMRSIVSELAGLSAMEYVQKLTDEEL 240
QY 180 SAGELSDQLNCGNSVTRQKMOOLEQMLTALDOMRSIVSELAGLSAMEYVQKLTDEEL 239
DB 180 SAGELSDQLNCGNSVTRQKMOOLEQMLTALDOMRSIVSELAGLSAMEYVQKLTDEEL 239
QY 239 ELADMKRRPELIACIGPPNICDLRENMWITSLAESQLTRQOIKKLEELQOKVSYKGDPI 298
DB 239 ELADMKRRPELIACIGPPNICDLRENMWITSLAESQLTRQOIKKLEELQOKVSYKGDPI 298
QY 240 ELADMKRRPELIACIGPPNICDLRENMWITSLAESQLTRQOIKKLEELQOKVSYKGDPI 299
DB 240 ELADMKRRPELIACIGPPNICDLRENMWITSLAESQLTRQOIKKLEELQOKVSYKGDPI 299
QY 299 VOHRPLLEERIVELFRLNLSAFVVEROPCMHPRPLVITKGVQFTTKVRLLVKFPPEL 358
DB 299 VOHRPLLEERIVELFRLNLSAFVVEROPCMHPRPLVITKGVQFTTKVRLLVKFPPEL 358
QY 300 IQHRPALLEERIVELFRLNLSAFVVEROPCMHPRPLVITKGVQFTTKVRLLVKFPPEL 359
DB 300 IQHRPALLEERIVELFRLNLSAFVVEROPCMHPRPLVITKGVQFTTKVRLLVKFPPEL 359
QY 359 NYOLKIVCIDKDSGDVAALRGSRRKFNILGTNTKVMNEESNNGSLSAEFKHLTLREQRC 418
DB 359 NYOLKIVCIDKDSGDVAALRGSRRKFNILGTNTKVMNEESNNGSLSAEFKHLTLREQRC 418
QY 360 NYOLKIVCIDKDSGDVAALRGSRRKFNILGTNTKVMNEESNNGSLSAEFKHLTLREQRC 419
DB 360 NYOLKIVCIDKDSGDVAALRGSRRKFNILGTNTKVMNEESNNGSLSAEFKHLTLREQRC 419
QY 419 GGRANCDASLVTEBELHLTFETEVYHOGKIDLETHSLRPVVVISNICOMPAASILMY 478
DB 419 GGRANCDASLVTEBELHLTFETEVYHOGKIDLETHSLRPVVVISNICOMPAASILMY 478
QY 420 GNGRPTNSDASLVTEBELHLTFETEVYHOGKIDLETHSLRPVVVISNICOMPAASILMY 479
DB 420 GNGRPTNSDASLVTEBELHLTFETEVYHOGKIDLETHSLRPVVVISNICOMPAASILMY 479
QY 479 WYNMLTRNPNKVNFFETKPPIGTMDQVAEVLWQFSSTTRKGLSIEQLTLAEKLLRPGVNS 538
DB 479 WYNMLTRNPNKVNFFETKPPIGTMDQVAEVLWQFSSTTRKGLSIEQLTLAEKLLRPGVNS 538
QY 480 WYNMLTRNPNKVNFFETKPPIGTMDQVAEVLWQFSSTTRKGLSIEQLTLAEKLLRPGVNS 539
DB 480 WYNMLTRNPNKVNFFETKPPIGTMDQVAEVLWQFSSTTRKGLSIEQLTLAEKLLRPGVNS 539
QY 539 YSGCOITWAKFCKENNAAGGFSFWWLNDIIDLKYYILALMNGYIMGFISKEERAIL 598
DB 539 YSGCOITWAKFCKENNAAGGFSFWWLNDIIDLKYYILALMNGYIMGFISKEERAIL 598
QY 540 YSGCOITWAKFCKENNAAGGFSFWWLNDIIDLKYYILALMNGYIMGFISKEERAIL 599
DB 540 YSGCOITWAKFCKENNAAGGFSFWWLNDIIDLKYYILALMNGYIMGFISKEERAIL 599

OY	599	STKPPGCTFLAFSESSKEGCVTFPMWVKDIDSGKTQIOISVEPYTKQOLNNNSFAEIIIMGYK	658
Db	600	SPKPPGCTFLAFSESSKEGCVTFPMWVKDIDSGKTQIOISVEPYTKQOLNNNSFAEIIIMGYK	659
OY	659	IMDANILVSPLVLYLPDIPKEAFAFGYCRPESQEHPEA--DPGSAA--PYLTKFKFCVY	714
Db	660	IMDANILVSPLVLYLPDIPKEAFAFGYCRPESQEHPEA--DPGSA--PYLTKFKFCVY	717
OY	715	P-----TTCSNFTIDL--PMSPRILDSLMQFUNGNEGAEPESAGQFE	753
Db	718	PCPSVFMDPPOSELLGNGFCPTGNSGNTSDLPMSPPRLTSLM----HNEAAEAANP--CPLE	772
OY	754	SLTFMDLITSECATSPM	770
Db	773	SLTJLDMELLSDDHA--SPM	788

```

RESULT      5
OI3133
ID      OI3133      PRELIMINARY;      PRT:      767 AA.
AC      OI3133;
DT      01-JUL-1997 (TEMBLrel. 04, Created)
DT      01-JUL-1997 (TEMBLrel. 04, Last sequence update)
DT      01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE      STAT3.
CN      HRTSTAT3.
OS      Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC      Procladanlopherygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX      NCBI_TaxID=8022;
RN      111
RP      SEQUENCE FROM N.A.
RA      Johnson M.C., Mourich D. V., Leong J.C.;
RL      Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR      EMBL; U60333; AAB60926.1; .
DR      HSSP; P42227; 1BCL
DR      InterPro; IPR000980; SH2.
DR      InterPro; IPR001217; STAT.
DR      Pfam; PF00017; SH2; 1.
DR      Pfam; PF01017; STAT; 1.
SQ      SEQUENCE 767 AA: 87816 MW: 927740C74C3798 CRC64:

```

Query Match	83.78	Score 3373.5	DB 13	Length 767
Best Local Similarity	84.48	Pred. No. 2.3e-220		
Matches 656	Conservative 53	Mismatches 51	Indels 17	Gaps 9

QY	1	MAONMLOOLDPRTYKJCHOLJYSPFPPELJYKOPFLAPMIESODMAYAAKSKEHATLYVHNL	60
Db	1	MAONMLOOLDPRTYKJCHOLJYSPFPPELJYKOPFLAPMIESODMAYAAKSKEHATLYVHNL	60
QY	61	LGEDIDOOYSRFLQESNVLYOHNLRIKOFLOSRYLEKPMETIARIVARCLWEEESFLIOTAA	120
Db	61	LGEDIDOOYSRFLQESNVLYOHNLRIKOFLOSRYLEKPMETIARIVARCLWEEESFLIOTAT	120
QY	121	TAOOGCGANHPHTAAVYVEKOOOMEOLHODVRRKVODLECKMKVENVLDDI'DENRYKTLK	180
Db	121	TAAD-GRASHPSGIVYKEKOILEHNLODRIKRVYODMECKMKLENVLDDIFDRHYTKL	179
QY	181	SOGDM-ODLNGNNO-SVTRÖKMOOLEBOMLTALDOMRISIVSELAGLLSAMEYOVKLTLE	238
Db	180	SOGELSDOMNGNSQAAARÖKMSOLEOMLSALDOLRROYTEMAGLLSAMPFOVKNLTTD	239
QY	239	ELADMKRRPEIACIGPBNICLDRLENNITSLAESOJOTROUKIKLELOOKSYKCDPI	298
Db	240	ELADMKRRQOJACIOGPXKICLDRLENNITSLGELIOLQIROKIKLELOOKSYKCDPI	299
QY	299	VOHPRMLEBRIVELFRNLKMSAFYVEROPCMHPRDPLVYKTVGFYTKVALLKVPFL	356
Db	300	IÖHRALEKEIYVDFLRNLKMSAFYVEROPCMHPRDPLVYKTVGFYTKVALLKVPFL	359

OY	359	NYO/LKIKVCJDKOSDVAALGSKRENTLGTNTEVMMSESNNSQSAEKKHLLTPEORC	418
		: : : : : : : : : : : : : : : : : : : :	
Dd	360	NYO/LKIKVI DKESGDVAAIKCSRKFNLLGNTKVMNEESNNQSIAEFKHLLTRKÖRC	419
OY	419	GNGGRANDASILVTVEELHLITFEFEVYHGOKLIDLETHSLPVAVISNICOMPNAVASIL	478
		: : : : : : : : : : : : : : : : : : : :	
Dd	420	GNGGTTNSDASILTVEELHLITFEFEVYHGOKLIDLETHSLPVAVISNICOMPRNASIL	479
OY	479	WYNMLTNPNKNVNFPTKPPJGTMDOVAEVLSMOFSSTTKRGLSTIEOLTTIAEKLLIGVN	538
		: : : : : : : : : : : : : : : : : : : :	
Dd	480	WYNMLTNPNKNVNFPTKPVGTMDOVAEVLSMOFSSTTKRGLTTEOLTTLAEKLLIGVN	539
OY	539	YSGCOITMAKCKEKMACKGSFWMYLDNIITDYKKYILAAMNGYIMGFISNERERAIL	598
		: : : : : : : : : : : : : : : : : : : :	
Dd	540	YSGCOITMAKCKEKMACKGSFWMYLDNIITDYKKYILAAMNGCYILGFISSERERAIL	599
OY	599	SIRPKPOTFLIPIRESKKGCGYTFMWEKDISGKTUOSVEPYTRKOOLNNMSFAEIIINGYK	658
		: : : : : : : : : : : : : : : : : : : :	
Dd	600	SKRPVYTELIFRIESSKEGGYTFTWEVDISGKTUHSVEPYTRKOOLANSIFAKIIMGYK	659
OY	659	IMDATNIIIVSPVLVLYPYDPKPEEARGAKCRESDOFHPHADPGSAA---PYLKTFPICVT	714
		: : : : : : : : : : : : : : : : : : : :	
Dd	660	IMDATNIIIVSPVLVLYFPKIPKEDAGNYCRDEAA--PRAFIGNPCITIOPLXKTFICVT	717
OY	715	PYTSCNTIDLPHSRRTLDLMOGFONNEGAPPSAGGOFSULTDHDLTSBCATSPM	770
		: : : : : : : : : : : : : : : : : : : :	
Dd	718	PYNSGNTSULPFMSRRTLNNISLT---HNKAENNP---GRLDLSLTDLMELSISVA-SIPM	767

```

RESULT      6
O9QXKO      PRELIMINARY;      PRT;      749 AA.
AC      O9QXKO;
DT      01-MAY-2000 (TREMBLrel. 13, Created)
DT      01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT      01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE      SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 1.
GN      STAT1.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxId=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Chen G., Newgard C.B.;
RT      "Expression of SPAT1 in INS-1 cells.";
RL      Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF205604; AAF20200.1; -.
DR      HSSP; P42224; 1BF5.
DR      InterPro; IPR000980; SH2.
DR      InterPro; IPR001217; STAT.
DR      Pfam; PF00017; SH2; 1.
DR      Pfam; PF001017; STAT; 1.
DR      SMART; SM00252; SH2; 1.
DR      PROSITE; PS50001; SH2; 1.
SQ      SEQUENCE 749 AA: 87234 MW: D37C634215DEN355 CRC64;

```

Query Match	50.68;	Score 2038;	DB 11;	Length 749;
Best Local Similarity	51.88;	Pred. No. 7e-130;		
Matches 396; Conservative	145;	Mismatches 195;	Indels 28;	Gaps 11;

QY	1	MAONMNOLOJDRYXJKOJLHOLYSIPFEPMELKOPFLAPWIESODMVAASKESHATYHPNL	60
Db	1	MSOMFELQOJDSKFLQVQHOLYDDBSEPMETKOYLAQWLEKDMEMAAWYDSFATIRHDL	60
QY	61	LCEDIOOYSRFLQESNVLYOHNLRIRIKOFLQDSRYLEKPMELARIVARCLMEBSHLLQTA	120
Db	61	LSOLDIOYSRFLQESNVLYOHNLRIRIKSRKSRNLQDNFQDEPVOXSMITHCLKEEKLLENAC	120
QY	121	--TAAGOGGGOAHNPAAVAVTEKQOMJEBHLODYVRKRVODLJOKMKVVENJLODDDFDENYKT	178
Db	121	RFNQAOQONJON---TYMLDKQKLSKVRNVKQDVWCISEHETKLTEDJLODVEDFCFT	176

Qy	179	LSKSGDMODLNGNNSYSTROKMOQLEOMLTALDQMRSYSELAGLISAMEYOKTLTDE	238
Dp	177	--SONRESEANGVAKSPOKQBOLLHMFMLMDKKREITIHKIRELLNSIELTONTLIND	234
Qy	239	ELAAMKRREPELACIGCPFNICIDBLEMWTISLAEQSLOTROQIKLEEDLOOKSYKGDPI	298
Dp	235	ELVEEMKRROQSACIGCPNMACLDQJSMFTIVASLDQVROQLKKELEEDQKFTYEGDP	294
Qy	239	VOHPRMIEERLEVELFRNLKMSAFVNERQCPMHPDRPLVIAKTVGTVOFTTVRYLLKFP	358
Dp	295	TKNRQVLSDRFVFLYEQOLIOSSFVERQPCPHRPORPVLKGTVOFTVRLLYKQLEL	354
Qy	359	NYOLKIKVCIDKDSGDVAALBGRSKFNILGNTKVMNMEESNNNGSLSAEKKHLTPREOR	418
Dp	355	NYNLKVKVSPKDVNENKNTVKGFEKFNILGHTKVMNMEESTNGSLAAEPRHLOLQKOK	413
Qy	419	GNGRANCDAASLYTEBELHLITFETVEYHOGKLIDLETHSLPVVVISNICOMPMAASIL	478
Dp	414	--NAGRNNEGPLITHEELHLSFETOLCQPLVYIDLETTSLPVVVISNVSOLPSGASIL	472
Qy	479	WYNMLTNPKVWNEFTKPPICPTQWQVAEVLISWQSSPTTKGLSIEOLTTLAEKLLGCVN	538
Dp	473	WYNMLYTERRLSFFLNPCCPMWSQBLEVLSWQSSVYTKRGLNDQJLSMLGKLLGNAG	532
Qy	539	YSGCOITWAKFCCKNNMAGKGFSPVWMLDNTIIDLVKKYILALWNGYIMGFSKERERAIL	598
Dp	533	PDGL-IPMTBRCKENINDKNFSEFWPIDTILIELKHKLLCLMNDQCIMGFIKREERALL	591
Qy	599	STKPRFTLLFSESSKEGGVTFWWEKDIS-CKTQIOSVEPYKQOOLNMNFAEITMGV	657
Dp	592	KDQGPFTLLRFSSSNEGATFTFWERSONGSGRPHAEPTIKRLSAVTPDIIRNY	651
Qy	658	KIMDATNLVSPLYLYPIDIPKEFAECK-YCRP-ESQEHPRADGSAAPYLKTKFCVT-	714
Dp	652	KVMAENIPEMLPYLYLPNIDKHAFCYKYSRPEAEPEMELDDPKRTGYIKELLISVSE	711
Qy	715	--PTTCSNTID-LPMSPRTDLSLQFNNNGGAPESAGGQFESL	755
Dp	712	VHPSRLSTENLRLPMSPDEPDEMSKI-----VGSPEDSM	745
RESULT	7		
Q9D323			
ID	Q9D323	PRELIMINARY;	PRT; 749 AA.
AC	Q9D323;		
DT	01-JUN-2001 (TREMBlrel. 17, Created)		
DT	01-JUN-2001 (TREMBlrel. 17, Last sequence update)		
DT	01-JUN-2001 (TREMBlrel. 17, Last annotation update)		
DE	SIGNAL PRODUCER AND ACTIVATOR OF TRANSCRIPTION I.		
GN	STAT1.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
RN	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=COLON;		
RX	MEDLINE=2108560; PubMed=11217851;		
RA	Kawai J., Shinaaga A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Akaiwa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,		
RA	Alizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.		
RA	Saito T., Okazaki Y., Gotohori T., Bono H., Kasudawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,		
RA	Knehl P., Lewis S., Matsuo Y., Nikido I., Passole G., Quackenbush J.,		
RA	Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,		
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
RA	Brownstein M.J., Bolt C., Fletcher C., Fujita M., Gardoloi M.,		
RA	Gustincich S., Hill D., Hofman M., Hume D.A., Kamiya M., Lee N.H.,		
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,		
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		

RA Sasakaki H., Sato K., Schoenbach C., Seya T., Shikata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.-H., Weltz C., Whiteaker C., Wilming L.,
RA Wyszynski-Boriss A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsunk S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.",
RL Nature 409:685-690(2001).
DR EMBL: AK018544; BAB31265.1; -.
DR MGD: MGI:103063; Scat1.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2: 1.
DR Pfam: PF01017; STAT: 1.
DR SMART: SM00252; SH2: 1.
DR PROSITE: PSS0001; SH2: 1.
SQ SEQUENCE 749 AA; 87280 MW; F7D40AC9D9ED7C99 CRC64;

Query Match	50.38;	Score 2026;	DB 11;	Length 749;
Best Local Similarity	51.78;	Pred. No. 4.6e-129;		
Matches 395;	Conservative 144;	Mismatches 197;	Indels 28;	Gaps 11;

```
OY      1 MAOMNOLOOLDTRYLKOJHOLYSDFPMELKOPAPWISQDMAYAAASKESATTLVFENI. 60  
       |::| ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB      1 MSQMFFLEUHLISKFLBGOVHQLYDSDSPMEIKOYLAWMLBKQUMFIMAYDVSPATIRFHDL. 60
```

```

QY      61 LGELDQYSRRLQESNVLQYHNLRIRKQELQSRKLEKPMELARIVAKGLWEESRLQATA 120
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      61 LSQLDDQYSRFSLENFLLQHNLRKSKKNLQDNFQEDPYQMSMLYNCLKEEKKILPNAU 120

```

```

Qy      121 -TAAQGGGCAANHPTAAVITEKQOQMLEQHLQDVKKRVQDTEQKMKVNIJQDFDFNKT 176
      111:1 1:1111:1:111:1 111:1:1 111:11 11
Db      121 RFNQOQEGNTQN----TWMLDKQKELDSKVRNKKVDQVMCIQEQIKTLEELQDEYDFCKKT 176

```

Db 177 --SQNREGGANGVAKSDQKQEDLLHKNFLMDPKRKELIHKIRELINSIELTQNTLND 234

235 ELVEKKRQDSACIGSPNACIDQLOSMTVAETLQOIROQLKEELEBKFTYEPPDI 294

Db 295 TKNNQVLSDRFLFLFOQLIQSSFVERQPCMPHPQRLVLKKTGVQFTVKRLVLVKLOEL 354

```

410 CACCGAAGCACT TATGTT ATT TTTTCTTTCCT TCTT CTTTCTTCCTGCAATGACCT 478
Db 355 NNNLKVVSFDKDVNEKNTVAGFRKFNILGTHKVMNMEESTNGSLAAEFRIQLKCEK - 413

```

```

Db      414  -MAGNRTNEGPLVTEELHSLSFETOLCQPGVLIDLETTSLPVVVISNVSQLPSCWASIL 472
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Ov      479  WYNNM.TNNPKNVNFETKPP ICTMDOVAEVI SWQFSSTTKRGISFIOI TTTIAEKI LGGVNV 518

```

[illegible]

Db 533 PDGL-IPWTFCKENINDKNFSFWPMIDTILELKKHLLCLWNDCIMGFI SKERERALL 591D

Qy 599 STKPDGFTLLRFSESKEGGVFTTWEMDIS-CKTQIQSVEPYTKOOLNNSFAEIIIMGY 6577

Db 592 KDDPGTFLRFSESSREGAITFTWVERSQNGCEPDPHAVEPYTKKLSAVTFDILIRNY 651

Qy 658 KIMDATNIIIVSPLVLYLPDIPKEEAFGK-YCRP-ESQEHPEADPGSAPYIKTKTFCVT- 714

Db 652 KMAAENIPENPLKLYLPNIDKDHGKYYSRKPEAPEMELDDPKRTGYIKTELISVSE 711

D5 712 VHPRLQTTDNLPRMSPEEFDEMSRI-----VSPFDSM 745

RESULT 8
ID 099K94 PRELIMINARY; PRT: 712 AA.
AC 099K94:
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC004808; AAH04808.1; -;
SO SEQUENCE 712 AA; 83106 MW; D62081709638CDA5 CRC64;

Query Match 49.8%; Score 2006; DB 11; Length 712;
Best Local Similarity 53.5%; Pred. No. 9,76-128;
Matches 385; Conservative 137; Mismatches 183; Indels 14; Gaps 8;

QY 1 MAONNOLOLDTRYLKOHLYSDTFPMELROFLAPWIESODMAYAAKESHATLVFNHL 60
DB 1 MSOMPELOOLSKFLEOVHLYDSFPMELROYLAOMLEKODMEHAAVDSQATVRFHDL 60
QY 61 LGEIDDOYSRFLQESNVLYOHNLRKIKOFLQSRYLEKPMETARIYARCLWESSRLLOTAA 120
DB 61 LSQDDDOYSRFLQESNVLYOHNLRKIKOFLQSRYLEKPMETARIYARCLWESSRLLOTAA 120
QY 121 --TAAGCGGQANHPPTAAVTEKQOMLEOHLODYKRRVODLEOKKAVVENLODDPFPNKT 178
DB 121 RENOAOEONIGN---TYMLDKOKELDSKVRNVQVQVNCIDEOIKTELEDEYDFKCKT 176
QY 179 LKSGODMODLNNOSVTRKMOOLEOMUTALDOMRSTVSELATLSAMEYVQCTLDE 238
DB 177 ---SQNRBEANGVAKSDOKOEOILLKMFMLDMKREKELIIRKRELLNSIELTQWTLND 234
QY 239 ELADKKRRPELACIGGPPNICLDRLENNITSLAESOLQTRQOIKLELOOKVSYKGDPI 298
DB 235 ELVEKKRRQOSACIGGPPNACLOQOSWFTTAEFTLOOIRQOLKLELEOKFTEYEPDI 294
QY 299 VOHRPMLEERIVELPRNLMSAFYVEROPCMHPRDPLVKTGYOFTTVKVLVKKFPEL 358
DB 295 TKNKOVLSDRFTFLFOQLIQSFVEROPCMPTHPORPLVLTGVOFTVKKLVLKLOEL 354
QY 359 NYOLKIKVCIDKSGDVAALRGSRKFNILGTNTKYNMNEESNGSLSAEFKHLTREORC 418
DB 355 NYNLKVKVSFKDVNEKMTVKGFRFNILGTHTKYNMNEESTNGSLAEFRHLQJKBOK 413
QY 419 GNGGANDASLIYTEELHLITFEVEYHOGKIDLETHSLPVPVVISNICOMPANASITL 478
DB 414 -NAGRTNEGLIVTEELHLSLFTFOLCOPGLVIDLETSLPVPVVISNVSQASITL 472
QY 479 WYNMUTNNPKVNFPTKPIGTMDQVAEVLWSQFSSTTKRGSLTDLTTLAKELGLPGVN 538
DB 473 WYNMUTNNPKVNFPTKPIGTMDQVAEVLWSQFSSTTKRGSLTDLTTLAKELGLPGVN 532
QY 539 YSGCOITAAKCKENMAKGGFSFWWLDNIIDLVKYYILALMNEGYIMGFISKEERAIL 598
DB 533 PDGL-IPTTRCKENINKNFSFPMIDTLEIKKHLCLMNDCCIMGFISKEERAIL 591
QY 599 STKPGGTFLRFSSKGGVTFWEKDIS-GKTOIOSVEPYTRKOOLNNMSFAITINGY 657
DB 592 KDOOGTFLRFSSKGGVTFWEKDIS-GKTOIOSVEPYTRKOOLNNMSFAITINGY 651
QY 658 KIMDATNLVSPVLYLDIPKEAFGR-YCRP-ESQEHPEADPGSAAPYLKTFKTCVT 714
DB 652 KYMAENIENPDLKYLINIDHAFGKYYSRPEKAPERMELDDPKRTGYIKTELISVS 710

RESULT 9
ID 013131 PRELIMINARY; PRT: 754 AA.
AC 013131:
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE SIGNAL TRANSDUCER/ACTIVATOR OF TRANSCRIPTION STAT1.
GN RBTSTAT1-1.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RA Johnson M.C., Mourich D.V., Leong J.C.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U60331; AAB60924.1; -;
DR HSSP: P42224; 1BF5.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SO SEQUENCE 754 AA; 86872 MW; 9C88B812B0DC6777 CRC64;

Query Match 47.2%; Score 1901; DB 13; Length 754;
Best Local Similarity 51.4%; Pred. No. 1,4e-120;
Matches 380; Conservative 140; Mismatches 200; Indels 20; Gaps 12;

QY 1 MAONNOLOLDTRYLKOHLYSDTFPMELROFLAPWIESODMAYAAKESHATLVFNHL 60
DB 1 MSOMPELOOLSKFLEOVHLYDSFPMELROYLAOMLEKODMEHAAVDSQATVRFHDL 60
QY 61 LGEIDDOYSRFLQESNVLYOHNLRKIKOFLQSRYLEKPMETARIYARCLWESSRLLOTAA 120
DB 61 LSQDDDOYSRFLQESNVLYOHNLRKIKOFLQSRYLEKPMETARIYARCLWESSRLLOTAA 118
QY 121 --TAAGCGGQANHPPTAAVTEKQOMLEOHLODYKRRVODLEOKKAVVENLODDPFPNKT 180
DB 121 RENOAOEONIGN---TYMLDKOKELDSKVRNVQVQVNCIDEOIKTELEDEYDFKCKT 176
QY 179 LKSGODMODLNNOSVTRKMOOLEOMUTALDOMRSTVSELATLSAMEYVQCTLDE 238
DB 177 ---SQNRBEANGVAKSDOKOEOILLKMFMLDMKREKELIIRKRELLNSIELTQWTLND 234
QY 239 ELADKKRRPELACIGGPPNICLDRLENNITSLAESOLQTRQOIKLELOOKVSYKGDPI 300
DB 235 ELVEKKRRQOSACIGGPPNACLOQOSWFTTAEFTLOOIRQOLKLELEOKFTEYEPDI 294
QY 299 VOHRPMLEERIVELPRNLMSAFYVEROPCMHPRDPLVKTGYOFTTVKVLVKKFPEL 360
DB 295 TKNKOVLSDRFTFLFOQLIQSFVEROPCMPTHPORPLVLTGVOFTVKKLVLKLOEL 354
QY 361 GNGGANDASLIYTEELHLITFEVEYHOGKIDLETHSLPVPVVISNICOMPANASITL 420
DB 361 GNGGANDASLIYTEELHLITFEVEYHOGKIDLETHSLPVPVVISNICOMPANASITL 409
QY 421 GGRANDASLIYTEELHLITFEVEYHOGKIDLETHSLPVPVVISNICOMPANASITL 480
DB 421 GGRANDASLIYTEELHLITFEVEYHOGKIDLETHSLPVPVVISNICOMPANASITL 468
QY 481 WYNMUTNNPKVNFPTKPIGTMDQVAEVLWSQFSSTTKRGSLTDLTTLAKELGLPGVN 538
DB 481 WYNMUTNNPKVNFPTKPIGTMDQVAEVLWSQFSSTTKRGSLTDLTTLAKELGLPGVN 528
QY 539 YSGCOITAAKCKENMAKGGFSFWWLDNIIDLVKYYILALMNEGYIMGFISKEERAIL 598
DB 539 YSGCOITAAKCKENMAKGGFSFWWLDNIIDLVKYYILALMNEGYIMGFISKEERAIL 586

```

Oy 599 STKPGCTFLLFSSSKRGCGVTFTHWVEKDISCKTQIOISVEPYTKQOLNNNSPFAEIIIMGYK 659
Db 567 TGKCGTFLLFSSSSRGCATFTTWVEHDQYNKVFHFAVEPYTKKELSAVS.LPDIIRTK 646
Oy 659 IMDATNILVSLVLYPDIPKEAFGKYYC--RPESQEHPEADPGSAPYLKTFICVPT 716
Db 647 VMAENIPENLRFRELYPDIPKDKSFGKYYARASPAESEPMDEGSSSTGYMKTETLSVSEV 706
Oy 717 TCSNTID--LPMSPRTDLSL 734
Db 707 HPSRLDDNNMPMSPDVPGEL 726

RESULT 10
013132
AC 013132: PRELIMINARY: PRT: 754 AA.
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE STAT1 -2.
GN R8TSTAT1-2.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RA Johnson M.C., Mourich D.V., Leong J.C.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U60332; AAB60925.1; -.
DR HSSP: P42224; 1BF5.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; SPAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SO SEQUENCE 754 AA; 87147 MW; 340B9645EA040142 CRC64;

```

[illegible]

Db	355	QJXVALLDQ---	-DVT	EKKGFRRKFNILGNTN	VMMNNEE-	NQSLAAEFRRHLDLQKEDQVA-	409
Oy	421	GGRANCDASLIVTEELHILITTE	EYVYHOC	LKIDLETHSLPVYVINSI	COMPANASILMY		480
Db	410	GMRNT-EGCLITTEELHIC	IFESSELNOS	LEKLELFTISLP	IVIVINSVQSLPGSMASILMY		468
Oy	481	NNLTNNPKVNFETPRPICTMD	VAEVLMSQSSPTTKRCLST	IEOLITLAEKLLGPRV-	N	538	
Db	469	NNLTSEPKRLKFEFLSPRAAS	MCQLESEVLMSQSSVYTKR	CLNEQLMLADKLLGQAKQ	NRN	528	
Oy	539	YSGCOITMAKFC	KENMACKGFSFWYLDNI	IDLVKYILALN	NEGYIMGFISKERERAIL	598	
Db	529	PEGL-IPWTKFEK	-SLSEKSPFPMI	IEAILDOLIKRHLSTL	NNDDCILLGFVKEKEREKML	586	
Oy	599	SRKPGCTFLRSESSKEG	YTFWVENDISGKTOLQSEV	PEPTKCOLNMSEAEI	IMCYK	658	
Db	587	TOKCCTFLRSESSRDOALIT	FWVHDLXKPVFHA	EPYTKKELSAVSL	LPDILIRYK	646	
Oy	659	IMDATNIIVSP	LYVLYPDLPKEAGK	KYC-RPEOSEH	PEADPGSAAPYLKTKFCVYPT	716	
Db	647	VMAENIPENPL	FLPYLPDIPDKSG	KYTYTRASEASEP	MDVGSSVGYMKTLLISVSEV	706	
Oy	717	TCSNTID--LPMSPRT	LDL 734				
Db	707	HPSRLQDNM	PMSPDVFGL 726				
RESULT 11							
	093598						
AC	093598		PRELIMINARY:			PRT: 749 AA.	
DT	01-NOV-1998		(TREMBLrel). 08.			Created)	
DT	01-NOV-1998		(TREMBLrel). 08.			Last sequence update)	
DT	01-JUN-2001		(TREMBLrel). 17.			Last annotation update)	
DE	TRANSCRIPT		FACTOR.				
GN	STAT1.						
OS	Brachydan: stat1 (zebrafish) (zebra danio).						
OC	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:						
OC	Actinopterygii: Neopterygii: Teleostei: Euteleostei: Ostariophysi:						
OC	Cypriniformes: Cyprinidae: Rasbora: Danio.						
RN	NCBI_TaxID: 7955;						
OX	[1]						
RP	SEQUENCE FROM N.A.						
RA	Oates A.C.						
RL	Thesis (1998), University of Melbourne, Australia.						
DR	EMBL: AJ005642; CA06676.1; ..						
DR	HSSP: P422		1NF5.				
DR	InterPro: IPR00980; SH2.						
DR	InterPro: IPR01217; STAT.						
DR	Pfam: PF01217; STAT.						
DR	Pfam: Pfam: STAT1.						
DR	SMART: Sm: SH2; 1.						
DR	PROSITE: PS50001; SH2; 1.						
SO	SEQUENCE 749 AA; 86856 MW; A84F66DEB2161137; Def64;						

Query Match	44.3%	Score 1786.5	DB 13	Length 749
Best Local Similarity	47.3%	Pred. No. 7.8e-113		
Matches 354	Conservative 159	Mismatches 202	Indels 33	Gaps 15
OY	1	MAOMNOLOOILDRYKOLKOLYSDPFPMPLROFLMWPWISQOMAAVAAS--KESHATLVFNH	58	
Db	1	MTQMLTFLQDLDRKFLBYDQQLYDDNFPMALIKRYLISWIESHMDHVAANYENSLATVFNH	60	
OY	59	NILGETDOOYSFLEDSNVLVOHNLRRIKOPLQSRYLEKPMETARIVARCLMEESRLQTH	118	
Db	61	DLTLDDHDSHFVWEKDFLNHOHNIIRKFKRNQNLFPDNPVSMAAMISHLNREEKKI--	118	
OY	119	AATAAOGGQAHNPPAAVUTTEKOYLNECLOLVQRKRYODLEKMKVVEVLQDDPPFNKT	178	
Db	119	-AIAOMSESPPRPNDODIVKNNRMDARVAKITKNKLVQDAEOTINLELDQDEYDFKNKT	177	
OY	179	LKSQGDMDLNG--NNQASTROKMOOLEQMLLALDOMRRSIYSELAGLSIAREYQKTLT	236	

```
DB 178 YOR---DELNGTISKEDLNREKL-MIHAMFVNLNDRIMATISOKEVNLVPEVICELI 233
OY 237 DEELADMKRRREIACIGCPNIIICLDRLNMTSLAESQLOTRQOIKKEELQKVSYKD 296
DB 234 SFLVEMKRRQOACIGCLTNVCLDQJONMFTSLGSLQSRQOKKLELEQKLSYSTD 293
OY 297 PIVQHRPMLEERVELFRNLKSAFVEROPCPMHDPRLVYIKTCVOFTTKRLLVKKP 356
DB 294 PTCGSKISLEERIMSNKLNVTNSLVEROPCPMTOLQRLPLVLTGVLTKRLKRLVKIP 353
OY 357 ELVYOLKIKYICITIKDGLVAAALRGSRKFNILGTNYVMNEESNNGSLSAEFKHLTLREQ 416
DB 354 EFNHTVAVKQOPKDI--VEKRGKGFHFNILGTSMKVMNNEESN--CMAEFPRLQLEK 409
OY 417 RCNGGRANCASLIYVEELHLITFEYVHQGLKIDLEFHSILPVVYISNICOPNMAAS 476
DB 410 KV--TGKRTSEGLVVEELHSLTFEALCWSGLVINFEFTSLPIVVISNVSOPLSGMAS 467
OY 477 ILVYNNMLTNPKNVNFTPKPIGMDQVAEVLISMOFSSTTKRGLSTEBOLTTLAELKIGP- 535
DB 468 IMYNNMLISEPKNLSFVTPPATWGOLEVLISMOFSSTTKRGLNPNQSLMGLKLLGPK 527
OY 536 GVNYSQGITWAKFCCKENMAGKGFSPVWMLDNIIDLVKYYILALNMGYIMGFISKER 595
DB 528 AASDFEAOIFMNRFCCKG-SEKNFTFWLMT EAVLDLIKRLHLSLMDGCIKMGFVTKDRTK 586
OY 596 AILSTRPGFTILRFESSKEGVFTWVEKDISGTOIOSVEPYTKOOLNMSFAETIM 655
DB 587 SLNNKAPGFTILRFSSNSNDGATTFSSWEHDANEDPOINSVEPYTKKEELSAVSLPILIR 646
OY 656 GYKIMDATNIIIVSPVLYLPDIJPKKEAFGK-YCRPESQEHF---EADPGSAAPYLKTK 709
DB 647 NYRVMAADNVNPNENLRLYLPQIPKDDAFKKYKTKPTDKQEMVDKRADEG---YISTP 702
OY 710 FICYT---PTTCSNTIDLPMSPTLDSL 734
DB 703 LIIISEMKPOEYENEL-MPMSPEVYGEEL 729

RESULT 12
O9DDJ8 PRELIMINARY: PRT: 394 AA.
AC O9DDJ8:
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE TRANSCRIPTION FACTOR STAT3 (FRAGMENT).
GN STAT3.
OS Brachydanio rerio (zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Rasbora; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Miyoshi K., Cui Y., Riedinger G., Robinson P., Lehoczky J., Zon L.,
RA Oka T., Dewar K., Hemmighausen L.;
RT "Structure of the Mouse Stat 3/5 Locus Evolution from Drosophila to
RT Zebrafish to Mouse."
RL Genomics 0:0-0(2001).
DR EMBL: AF322857; AAC42495.1; -.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
FT NON_TER 1
SQ SEQUENCE 394 AA: 44126 MW: 28DEDEFE716B50E70 CRC64;
```

Query Match 39.7%; Score 1597.5; DB 13; Length 394;

Best Local Similarity 82.6%; Pred. No. 2e-100; Matches 317; Conservative 13; Mismatches 21; Indels 33; Gaps 8;

```
OY 412 TLREQRCNGGRANCASLIYVEELHLITFEYVHQGLKIDLEFHSILPVVYISNICOMP 471
DB 1 TLREQRCNGGRANCASLIYVEELHLITFEYVHQGLKIDLEFHSILPVVYISNICOMP 60
OY 472 NAASTILMYMMLTNPKNVNFTPKPIGMDQVAEVLISMOFSSTTKRGLSTEBOLTTLAEL 531
DB 61 NAASTILMYMMLTNPKNVNFTPKPIGMDQVAEVLISMOFSSTTKRGLSTEBOLTTLAEL 120
OY 532 LAGPGVNSQGITWAKFCCKENMAGKGFSPVWMLDNIIDLVKYYILALNMGYIMGFISK 591
DB 121 LAGPGVNSQGITWAKFCCKENMAGKGFSPVWMLDNIIDLVKYYILALNMGYIMGFISK 180
OY 592 EERRALLSTRPGFTILRFESSKEGVFTWVEKDISGTOIOSVEPYTKOOLNMSFA 651
DB 181 EERRALLSTRPGFTILRFESSKEGVFTWVEKDISGTOIOSVEPYTKOOLNMSFA 240
OY 652 EIIIMGYKIMDATNIIIVSPVLYLPDIJPKKEAFGK-YCRPESQEHF---DYGSA-PLYK 707
DB 241 EIIIMGYKIMDATNIIIVSPVLYLPDIJPKKEAFGK-YCRPESQEHF---DYGSA-PLYK 298
OY 708 TKFLCYTP-----TTCSNTIDL-PMSPTLDSIMQFGNNGCAEP 746
DB 299 TKFLCYTPCPSVEMDPDSELGNGFPGTNSGTSLEFPMSPTLDSLM---HNEAFA 354
OY 747 SAGGOFESTLTFMDLTSECATSPM 770
DB 355 NP-GPLESLTLDMLSSDHA-SPM 376

RESULT 13
O9QXJ2 PRELIMINARY: PRT: 922 AA.
AC O9QXJ2:
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 2.
GN STAT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Farrar J.D., Smith J.D., Murphy T.L., Murphy K.M.;
RT "Sequence Divergence in the C-Termini of Murine and Human Stat2
RT Controls Species-Specific Stat4 Activation and Th1 Development Induced
RT by Type I Interferons."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF206162; AAF17096.1; -.
DR HSSP: P42224; 1BF5.
DR MGD: MGI:103039; Stat2.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SQ SEQUENCE 922 AA: 105365 MW: 61E73D3276D5354A CRC64;
```

Query Match 31.9%; Score 1287; DB 11; Length 922;
Best Local Similarity 39.3%; Pred. No. 7.3e-79; Matches 294; Conservative 140; Mismatches 259; Indels 56; Gaps 20;

```
OY 1 MAQNNLOLDTRYKLOHLYSDT-FPMELRQFLAPMIESQDMAYAA--SKESHATVE 57
DB 1 MAQNNLOLDTRYKLOHLYSDT-FPMELRQFLAPMIESQDMAYAA--SKESHATVE 60
```


<C Eukaryota: Metazoa Chordata; Craniata; Vertebrata: Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxId=8022;
RN [1]
RP SEQUENCE FROM N.A.
RA Johnson M.C., Mourich D.V., Leong J.C.;
RT "Rainbow Trout STAT Genes: Evolutionary Duplication Leads to STAT 5A
RT and the Acquisition of Mammalian Specific Function."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF147499; AAG14946.1; -
DR InterPro: IPR000980; SH2; -
DR InterPro: IPR001217; STAT.
DK Pfam: PF00017; SH2; 1.
DK Pfam: PF01017; STAT; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SQ SEQUENCE 784 AA: 89780 MW: 1C16DCFC7C83AF7C CRC64;

Query Match 21.7%; Score 873; DB 13; Length 784;
Best Local Similarity 30.4%; Pred. No. 6,6e-51;
Matches 248; Conservative 151; Mismatches 314; Indels 104; Gaps 29;

OY 1 MAOMNLOOLDPFRYLYKQLHQLYSDPPELROFLAPMESODMAYV---ASKESHATLY 56
DB 1 MAVMTQAOQLQGLDALHQMOSLYGONHPIEVRYHLSQMLEGQLMDVIDLENPOEEFKAKRI 60
OY 57 PNLGELIDOOYSRFLQESNVLYQHNLRRIKQFLOSRYLEKPMETARIVARCLMESRLL 116
DB 61 LDLSLQELONKKAHEGDEGFLKIKLGHVASQLKSTYDRCPLELVRCIKHILYTEQRLV 120
OY 117 QTAATAAO--GGOANHPTAAVTEKQMLEQHLQDYRKRVQDLEQMKVVENLQDDPDFN 175
DB 121 REASNSSSPVGGMDS-----MSQKYQIQAFAEELRLITDPTENDLRKIQHNQEVFIIO 175
OY 176 Y-KTLKSGQDMODL-----NGNNSVTRKMOOLEQMLT-----ALDMRSIVSELG 223
DB 176 YQESLRQIQQLSLTLPAPADQLREPTLSKRAYEAWLTREANTLQKRYDLAKHOK 235
OY 224 LLSAMEYVOKTLTDEELADMKRREPETACIGRPNICDLRLNMTSLAESQLQTRQOIKK 283
DB 236 SLQLRKQOTIILDELIMHKRROQLAGNGGPEEGDILIOSCEKLAETIMQNRQIRR 295
OY 284 LLELOOKVSKGDPPIQVQHPMLERIVELFRNLMSAFVERQPCPMHPDRPIVKTGV 343
DB 296 AEHLRQQLPIPG-PIEELLNDLNSTITDIIISALVSTFIIEKOP-----PQVLKTOT 346
OY 344 OFTKVRLVKEPELNYOL---KIKVICDKSDGVAAALGRKFN-----ILGTNTKVMN 395
DB 347 KEAATVRLVG--GKLNVHNPPOVKATIIISQOAKALLKNENMRNDSGEIILNNC--VME 404
OY 396 MEESNNGSLSAEFKHLTLREQCGNGRANCOASLIYTELHLITFETEVYHOG--LKID 453
DB 405 YHQV--TGLTISANFRNMSLK--RIKRSDRGAES--VTEKEFTILFESQFSVQGNELVQ 458
OY 454 LETHSLPVVVISNICMPAMASITLWYNMLTNPKVNFETKPIGTWDOVAEVLMSQFS 513
DB 459 VKTSLSPVVVIYHGSDNNATATVLDNAFA--EPGRVPTIV--PDKVLMPOLCEALNMKTK 516
OY 514 S--TTKRGISIEQLTLAEKLLGPV---NYSQOITWAKFCKENMAGKGFSPVWILDN 567
DB 517 AEVQSNRGLSEENLYVLAQKAFSSSINPDYRGMTMTWSQFNRESLPGRNFTFMQMFQD 576
OY 568 IIDLVKKYILALMNEGYINGFISKEREKAILSTKPPGTFLLRFSSSKEGVYFTWEKD 627
DB 577 VVELTKRHLKPHMNDALIGFVNKOAOADMLSKPNGTFLRFSQ--SEIGGITIAVVAEN 635
OY 628 IS--GKTOISVEPYTKOOLNMMSFAEILMGYKIMDATNILVPLVYLVPDIPKEEAFGR 685
DB 636 PKKPGERWVWMLMPTTKDFISRLAD-----RISDLNHL-----FLYPRDKRDEVFESK 685
OY 686 YCRPESQEHPEADPESAA--PYLTKFKICVTPPTCSNTIDLPMSPRTILDSLMQFGNNGEG 743

DB 686 YTP-----PLSKAVDGIVKPOIKOVVPEFTTTNPDPAANPTVMU-----HG 727
OY 744 AEPSAG-----GQFESIT-----FDMULTSECA 766
DB 728 ASPVSHPPNNGIYQPMDSILDADGDFOLDPTMDVA 764

Search completed: March 19, 2002, 15:25:44
Job time: 226 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 19, 2002, 15:19:38 ; Search time 23.78 Seconds

(Without alignments)
1187.213 Million cell updates/sec

Title: US-08-212-185-12

Perfect score: 4029
Sequence: 1 MAQWNLQQLDRLRYLKLQHLQ.....QFESLTFMDLTSECATSPM 770

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4010	99.5	770	1	STRA3_MOUSE
2	4007	99.5	770	1	STRA3_MOUSE
3	3977	98.7	770	1	STRA3_MOUSE
4	2017	50.1	750	1	STRA1_HUMAN
5	1984	49.2	749	1	STRA1_MOUSE
6	1798	44.6	748	1	STRA4_MOUSE
7	1780.5	44.2	749	1	STRA4_MOUSE
8	1293	32.1	851	1	STRA2_MOUSE
9	1271.5	31.6	923	1	STRA2_MOUSE
10	1264.5	31.4	864	1	STRA2_MOUSE
11	867	21.5	799	1	STRA2_MOUSE
12	865.5	21.5	786	1	STRA2_MOUSE
13	865.5	21.5	786	1	STRA2_MOUSE
14	862.5	21.4	786	1	STRA2_MOUSE
15	862.5	21.4	786	1	STRA2_MOUSE
16	859	21.3	787	1	STRA2_MOUSE
17	858.5	21.3	794	1	STRA2_MOUSE
18	858	21.3	793	1	STRA2_MOUSE
19	857	21.3	793	1	STRA2_MOUSE
20	855	21.2	787	1	STRA2_MOUSE
21	791.5	19.6	794	1	STRA2_MOUSE
22	582	14.4	837	1	STRA6_MOUSE
23	558.5	13.9	847	1	STRA6_MOUSE
24	481	11.9	761	1	STAT1_MOUSE
25	159	3.9	1085	1	STAT1_MOUSE
26	148	3.7	2663	1	STAT1_MOUSE
27	145	3.6	1427	1	STAT1_MOUSE
28	145	3.6	1433	1	STAT1_MOUSE
29	145	3.6	1433	1	STAT1_MOUSE
30	144.5	3.6	1959	1	STAT1_MOUSE
31	143	3.5	1087	1	STAT1_MOUSE
32	141	3.5	3911	1	STAT1_MOUSE
33	140.5	3.5	1312	1	STAT1_MOUSE

34	140.5	3.5	1526	1	MYO2_MOUSE	O9u516
35	140.5	3.5	2230	1	G0G4_MOUSE	O13439
36	140	3.5	1940	1	MYH3_MOUSE	P02565
37	140	3.5	3685	1	DMY_MOUSE	P11532
38	138.5	3.4	1939	1	MYH1_MOUSE	P12882
39	138	3.4	2411	1	MYO2_MOUSE	P05661
40	137.5	3.4	1818	1	MYH2_MOUSE	P75471
41	137	3.4	1978	1	MYH2_MOUSE	P10587
42	136	3.4	857	1	MYH2_MOUSE	P10587
43	135.5	3.4	1505	1	MYH2_MOUSE	P10587
44	135.5	3.4	1937	1	MYH2_MOUSE	P13335
45	135.5	3.4	2349	1	MYH2_MOUSE	P12270

ALIGNMENTS

RESULT	ID	STRA3_MOUSE	STANDARD:	PRT:	770 AA.
AC	P42227	STRA3_MOUSE	STANDARD:	PRT:	770 AA.
DT	01-NOV-1995 (Rel. 32, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	20-AUG-2001 (Rel. 40, Last annotation update)				
DE	SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 3 (ACUTE-PHASE RESPONSE FACTOR).				
DE	START3 OR ABRF.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 154-158; 181-185 AND 632-640.				
RC	STRAIN=BAIB/C; TISSUE=Liver;				
RX	MEDLINE=94208062; PubMed=7512451;				
RA	Akira S., Nishio Y., Inoue M., Wang X.J., Shi W., Matsusaka T.,				
RA	Yoshida K., Sudo T., Naruto M., Kishimoto T.;				
RT	"Molecular cloning of ABRF, a novel IFN-stimulated gene factor 3 p91-				
RT	related transcription factor involved in the gp130-mediated signaling				
RT	pathway.";				
RL	Cell 77:63-71(1994).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Thymus;				
RX	MEDLINE=94188718; PubMed=8140422;				
RA	Zhong Z., Wen Z., Darnell J.E. Jr.;				
RT	"Stat3: a Stat family member activated by tyrosine phosphorylation in				
RT	response to epidermal growth factor and interleukin-6.";				
RL	Science 264:95-98(1994).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Brain;				
RX	MEDLINE=95014185; PubMed=7523373;				
RA	Raz R., Durbin J.E., Levy D.E.;				
RT	"Acute phase response factor and additional members of the				
RT	interferon-stimulated gene factor 3 family integrate diverse signals				
RT	from cytokines, interferons, and growth factors.";				
RL	J. Biol. Chem. 269:24391-24395(1994).				
RN	[4]				
RP	SEQUENCE FROM N.A. (STAT3B).				
RC	STRAIN=BAIB/C, AND B6; TISSUE=Liver;				
RX	MEDLINE=96016116; PubMed=7568080;				
RA	Schaefer T.S., Sanders L.K., Nathans D.;				
RT	"Cooperative transcriptional activity of Jun and Stat3 beta, a short				
RT	form of Stat3.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 92:9097-9101(1995).				
RN	[5]				
RP	PHOSPHORYLATION OF SER-727, AND MUTAGENESIS.				
RX	MEDLINE=95354205; PubMed=7543024;				
RA	Wen Z., Zhong Z., Darnell J.E. Jr.;				
RT	"Maximal activation of transcription by Stat1 and Stat3 requires both				
RT	tyrosine and serine phosphorylation.";				
RL	Cell 82:241-250(1995).				


```

RL J. Biol. Chem. 270:29998-30006(1995).
CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE INTERLEUKIN-6
CC (IL-6)-RESPONSIVE ELEMENTS IDENTIFIED IN THE PROMOTERS OF VARIOUS
CC ACUTE-PHASE PROTEIN GENES.
CC -1- PATHWAY: INVOLVED IN THE Gp130-MEDIATED SIGNALING PATHWAY.
CC -1- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
CC MEMBER (AT LEAST STAT1) (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION (BY SIMILARITY).
CC -1- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IL-6, IL-11, CNTF,
CC LIF, CSF-1, EGF, PDGF, IFN-ALPHA AND OSM. SERINE PHOSPHORYLATION
CC IS IMPORTANT FOR THE FORMATION OF STABLE DNA-BINDING STAT3
CC HOMODIMERS AND MAXIMAL TRANSCRIPTIONAL ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR FMHI: X91810: CAA62920.1:
DR HSSP: P42224: 1BFS
DR InterPro: IPR000960: SH2
DR InterPro: IPR001217: STAT.
DR Pfam: PF00017: SH2; 1.
DR SMART: SM00252: SH2; 1.
DR PROSITE: PS00001: SH2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW Phosphorylation; SH2 domain.
FT MOD_RES 580 670 SH2.
FT DOMAIN 580 705 PHOSPHORYLATION (BY JAKS) (BY
FT MOD_RES 705 705 SIMILARITY).
FT MOD_RES 727 727 PHOSPHORYLATION (BY SIMILARITY).
FT FT SEQUENCE 770 AA: 86039 MW: D74A0C76954754ED CRC64;
SQ

```

```

Query Match          99.5%; Score 4007; DB 1; Length 770;
Best Local Similarity 99.4%; Pred. No. 1.3e-243;
Matches 765; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 MAQWNOLOQLDTRYLRQLHQLYSDFPMLRQFLAPWIESQDMAYAAKSKESHATVFNHL 60
DB 1 MAQWNOLOQLDTRYLRQLHQLYSDFPMLRQFLAPWIESQDMAYAAKSKESHATVFNHL 60
OY 1 LGELIDQYSEFLEDSNVLYOHNLRRIKOFLQSRLEKPMETARIVARCLMEESRLDTAA 120
DB 1 LGELIDQYSEFLEDSNVLYOHNLRRIKOFLQSRLEKPMETARIVARCLMEESRLDTAA 120
OY 121 TAAQGGGAQAHPTAAVYTEKQMLEOHLQDVRRVODLEQKMKVENVLQDDDFENYTKL 180
DB 121 TAAQGGGAQAHPTAAVYTEKQMLEOHLQDVRRVODLEQKMKVENVLQDDDFENYTKL 180
OY 121 TAAQGGGAQAHPTAAVYTEKQMLEOHLQDVRRVODLEQKMKVENVLQDDDFENYTKL 180
DB 121 TAAQGGGAQAHPTAAVYTEKQMLEOHLQDVRRVODLEQKMKVENVLQDDDFENYTKL 180
OY 181 SOGDMDLGNNSQVTRQKMQOLEOMLTALDQMRISVSELGLLSAMEYVQKLTLDDEL 240
DB 181 SOGDMDLGNNSQVTRQKMQOLEOMLTALDQMRISVSELGLLSAMEYVQKLTLDDEL 240
OY 181 SOGDMDLGNNSQVTRQKMQOLEOMLTALDQMRISVSELGLLSAMEYVQKLTLDDEL 240
DB 181 SOGDMDLGNNSQVTRQKMQOLEOMLTALDQMRISVSELGLLSAMEYVQKLTLDDEL 240
OY 241 ADMKRPETACIGPPNICIDRLNNMTSLAESOLOTRQOIKKLELQOKVSKGDPVYQ 300
DB 241 ADMKRPETACIGPPNICIDRLNNMTSLAESOLOTRQOIKKLELQOKVSKGDPVYQ 300
OY 301 HRPMLEBERIVLEFRNLKMSAFVVERQPCMPHMDRPLVITGVQFTTKVLLVKFPPLNY 360
DB 301 HRPMLEBERIVLEFRNLKMSAFVVERQPCMPHMDRPLVITGVQFTTKVLLVKFPPLNY 360
OY 361 QLKIKVICIKDSQDVAAKRSRKNLGLTNTKYMMNEESNGSLSAEFKHLTLREORCGN 420
DB 361 QLKIKVICIKDSQDVAAKRSRKNLGLTNTKYMMNEESNGSLSAEFKHLTLREORCGN 420

```

```

OY 421 GGRANDASLVITBEHLITFETEVYHUGLKIDLETHSLPVVVISNICQMPNMAASILMY 480
DB 421 GGRANDASLVITBEHLITFETEVYHUGLKIDLETHSLPVVVISNICQMPNMAASILMY 480
OY 481 NMLTNPNKVNPFTRKPPIGTDQVAEVLNMQPSSSTTKGSLTEQLTLAEKLLGCVNYS 540
DB 481 NMLTNPNKVNPFTRKPPIGTDQVAEVLNMQPSSSTTKGSLTEQLTLAEKLLGCVNYS 540
OY 541 GCOITMAFCEENNAKGFSEFWMLDNIIDLKRYKILALNMEGYMGFISKEREALIST 600
DB 541 GCOITMAFCEENNAKGFSEFWMLDNIIDLKRYKILALNMEGYMGFISKEREALIST 600
OY 601 KPGCTFLFRSESSKEGGVFTFWWEKDISGRVQIQSVEPYTKOOLNNNSFAELIMGYIM 660
DB 601 KPGCTFLFRSESSKEGGVFTFWWEKDISGRVQIQSVEPYTKOOLNNNSFAELIMGYIM 660
OY 661 DATNIVLSPLYLYPDIPKEAFKRCPEOEHPHEADPGSAAPLKRFICVPTTCSN 720
DB 661 DATNIVLSPLYLYPDIPKEAFKRCPEOEHPHEADPGSAAPLKRFICVPTTCSN 720
OY 721 TIDPMSPTLDSLMQFGNNGEAGPSAGQFESITFPMDLTSECATSPM 770
DB 721 TIDPMSPTLDSLMQFGNNGEAGPSAGQFESITFPMDLTSECATSPM 770

```

```

RESULT 3
STAT3_HUMAN STANDARD; PRT; 770 AA.
ID STAT3_HUMAN SH2;
AC P40763;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 3 (ACUTE-PHASE
DE RESPONSE FACTOR).
GN STAT3 OR ATRF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_Taxid 9606;
RN 1;
RP SEQUENCE: N. A.
RC TISSUE: placenta.
RA MEDLINE=94208062; Pubmed=7512451;
RA Akira S., Nishio Y., Inoue M., Wang X.-J., Wei S., Matsusaka T.,
RA Yoshida K., Suho T., Naruto M., Kishimoto T.;
RT "Molecular cloning of ATRF, a novel IFN-gamma-induced gene factor 3 p91-
RT related transcription factor involved in the IFN-gamma mediated signaling
RT pathway."
RL Cell 77:611-619 (1994).
RN 121;
RP PHOSPHORYLATION: ON SERINE.
RX MEDLINE=94208062; Pubmed=7701321;
RX Zhang X., Zhang J., Li H.-C., Schindler C., Chan-Kiang S.;
RT "Requirement of serine phosphorylation for formation of STAT-promoter
RT complexes."
RL Science 267:11990-11994 (1995).
CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE INTERLEUKIN-6
CC (IL-6)-RESPONSIVE ELEMENTS IDENTIFIED IN THE PROMOTERS OF VARIOUS
CC ACUTE-PHASE PROTEIN GENES.
CC -1- PATHWAY: INVOLVED IN THE Gp130-MEDIATED SIGNALING PATHWAY.
CC -1- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
CC MEMBER (AT LEAST STAT1).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION.
CC -1- TISSUE SPECIFICITY: HEART, BRAIN, PLACENTA, LUNG, LIVER, SKELETAL
CC MUSCLE, KIDNEY, AND PANCREAS.
CC -1- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IL-6, IL-11, CNTF,
CC LIF, CSF-1, EGF, PDGF, IFN-ALPHA AND OSM. SERINE PHOSPHORYLATION
CC IS IMPORTANT FOR THE FORMATION OF STABLE DNA-BINDING STAT3
CC HOMODIMERS AND MAXIMAL TRANSCRIPTIONAL ACTIVITY.
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: L29277; AAA58374.1; -
DR HSSP: P42224; 1BF5.
DR TRASNPA: T01493; -
DR MIM: 102582; -
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW Phosphorylation; SH2 domain.
FT DOMAIN 580 670 SH2.
FT MOD_RES 580 705 PHOSPHORYLATION (BY JAKS) (BY
FT MOD_RES 727 727 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 770 AA: 88052 MW: A3DCEE815B3B5360 CRC64:

Query Match 98.7%; Score 3977; DB 1; Length 770;
Best Local Similarity 98.6%; Pred. No. 9,8e-242;
Matches 759; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 MAONNOLOQDTRYLKOHLYSDFFPMELOFLAPMIESQDMAVAASKESATLVFNNL 60
DB 1 MAONNOLOQDTRYLKOHLYSDFFPMELOFLAPMIESQDMAVAASKESATLVFNNL 60
QY 61 LGEDIQOYSRFLQSSNVLVYOHNLRIKOFLOSRYLEKPMETARIYARCLMESSRLQTA 120
DB 61 LGEDIQOYSRFLQSSNVLVYOHNLRIKOFLOSRYLEKPMETARIYARCLMESSRLQTA 120
QY 121 TAAOOGGAGNHPPTAAVTEKQOMLEOHLYQDYRKRVODLEQMKVVENLODDFENYKTLK 180
DB 121 TAAOOGGAGNHPPTAAVTEKQOMLEOHLYQDYRKRVODLEQMKVVENLODDFENYKTLK 180
QY 181 SOGMODLNGNNOSTYRKMOOLEOMLTALDOMRSIVSELAGLISAMEYOKTLTDEEL 240
DB 181 SOGMODLNGNNOSTYRKMOOLEOMLTALDOMRSIVSELAGLISAMEYOKTLTDEEL 240
QY 241 ADMKRREIACIGPPNICLDRLNEMITSLAESOLQTRQOIKLELOQKYSYKGDPIVQ 300
DB 241 ADMKRREIACIGPPNICLDRLNEMITSLAESOLQTRQOIKLELOQKYSYKGDPIVQ 300
QY 301 HRPMLERIVELFRNLMSAFVVERQPCMPHPPRPVLIKGVQFTTKVRLVKEPELNY 360
DB 301 HRPMLERIVELFRNLMSAFVVERQPCMPHPPRPVLIKGVQFTTKVRLVKEPELNY 360
QY 361 OLKIKVCDIKDSGVAAALRGSRKFNILGTNTKVMNMESSNNGSLSAEKKHILTRPDRCGN 420
DB 361 OLKIKVCDIKDSGVAAALRGSRKFNILGTNTKVMNMESSNNGSLSAEKKHILTRPDRCGN 420
QY 421 GGRANCASLIVTEELHILTFETEVYHOGKIDLETSLPVAVSNICOMPANAFSTILMY 480
DB 421 GGRANCASLIVTEELHILTFETEVYHOGKIDLETSLPVAVSNICOMPANAFSTILMY 480
QY 481 NMLTNKRNVAFPTKPTGTMDOVAEVLVSWQSSSTTKRGLSIEOLTTLAEKLLGPGVNS 540
DB 481 NMLTNKRNVAFPTKPTGTMDOVAEVLVSWQSSSTTKRGLSIEOLTTLAEKLLGPGVNS 540
QY 541 GCOITTMKFCENNAGKGFPMWLDNITLVKYVYIALAMEGYIMGTSIKERERAILST 600
DB 541 GCOITTMKFCENNAGKGFPMWLDNITLVKYVYIALAMEGYIMGTSIKERERAILST 600
QY 601 KPPGTFLLRFSESSKEGVTFTWEKDISGKTQIOSVEPYTKOOLNNNSFAEIIIMGYKIM 660
DB 601 KPPGTFLLRFSESSKEGVTFTWEKDISGKTQIOSVEPYTKOOLNNNSFAEIIIMGYKIM 660

DB 601 KPPGTFLLRFSESSKEGVTFTWEKDISGKTQIOSVEPYTKOOLNNNSFAEIIIMGYKIM 660
QY 661 DATNLVSPFLVLYYDIPKEAFGKYCRPESQEHPEADGPSAPYLKTFICVFTPTCSN 720
DB 661 DATNLVSPFLVLYYDIPKEAFGKYCRPESQEHPEADGPSAPYLKTFICVFTPTCSN 720
QY 721 TIDLPMSPRTLDLMQFGNNGEAEPDSACGFESLTFMDLTSSECATSPM 770
DB 721 TIDLPMSPRTLDLMQFGNNGEAEPDSACGFESLTFMDLTSSECATSPM 770
RESULT 4
STAI_HUMAN STANDARD: PRT; 750 AA.
AC P42224;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 1-ALPHA/BETA
DE (TRANSCRIPTION FACTOR ISGF-3 COMPONENTS P91/P84).
GN STAT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 514-524; 654-660 AND 667-672.
RX MEDLINE=92366557; PubMed=1502203;
RA Schindler C., Fu X.-Y., Impirota T., Aebersold R., Darnell J.E. Jr.;
RT "Proteins of transcription factor ISGF-3: one gene encodes the 91 and
RL 84-kDa ISGF-3 proteins that are activated by interferon alpha";
RL Proc. Natl. Acad. Sci. U.S.A. 89:7836-7839(1992).
RN [2]
RP PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=95192056; PubMed=7885841;
RA Yan R., Qureshi S., Zhong Z., Wen Z., Darnell J.E. Jr.;
RT "The genomic structure of the STAT genes: multiple exons in
RL coincident sites in Stat1 and Stat2";
RL Nucleic Acids Res. 23:459-463(1995).
RN [3]
RP PHOSPHORYLATION OF TYR-701.
RX MEDLINE=95386533; PubMed=7657660;
RA Quella F.W., Thierfelder W., Wiltuhn B.A., Tang B., Cohen S.,
RL Thie J.N.;
RT "Phosphorylation and activation of the DNA binding activity of
RT purified Stat1 by the Janus protein-tyrosine kinases and the
RL epidermal growth factor receptor";
RL J. Biol. Chem. 270:20775-20780(1995).
RN [4]
RP PHOSPHORYLATION OF SER-727, AND MUTAGENESIS.
RX MEDLINE=95354205; PubMed=7543024;
RA Wen Z., Zhong Z., Darnell J.E. Jr.;
RT "Maximal activation of transcription by Stat1 and Stat3 requires both
RT tyrosine and serine phosphorylation";
RL Cell 82:241-250(1995).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 136-710.
RX MEDLINE=98292180; PubMed=9630226;
RA Chen X., Vinkemeier U., Zhao Y., Jeruzalmi D., Darnell J.E. Jr.,
RL Kurlyan J.;
RT "Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to
RT DNA";
RL Cell 93:827-839(1998).
CC -I- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE IFN-STIMULATED
CC RESPONSE ELEMENT (ISRE) AND TO THE GAS ELEMENT. THIS MULTIPROTEIN
CC TRANSCRIPTION FACTOR IS TERMED ISGF3.
CC -I- PATHWAY: IFN-SIGNALING PATHWAY
CC -I- SUBUNIT: IN RESPONSE TO IFN ALPHA/BETA, THREE SUBUNITS (STAT1-
CC ALPHA, STAT1-BETA, STAT2) OF ISGF3, BECOME PHOSPHORYLATED ON
CC TYROSINE, MIGRATE INTO THE NUCLEUS, AND ASSEMBLE INTO A COMPLEX
CC TOGETHER WITH ISGF3 GAMMA (P48), A DNA-BINDING PROTEIN THAT
CC SPECIFICALLY BINDS TO THE IFN-STIMULATED RESPONSE ELEMENT. IN
CC RESPONSE TO IFN GAMMA, STAT1 FORMS HOMODIMERS, THAT ALSO


```
CC TRANSCRIPTIONAL ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U06924: AAA19454.1: -.
DR HSSP: PA2224: 1BF5.
DR TRANSFAC: T01575: -.
DR MGD: MGI:103063: Stat1.
DR InterPro: IPR000980: SH2.
DR InterPro: IPR001217: STAT.
DR Pfam: PF00017: SH2: 1.
DR SMART: SM00252: SH2: 1.
DR PROSITE: PS50001: SH2: 1.
DR Transcription regulation; DNA-binding; Nuclear protein; SH2 domain;
KW Phosphorylation.
FT DOMAIN 573 670 SH2.
FT MOD_RES 701 701 PHOSPHORYLATION (BY JAKS) (BY
FT SIMILARITY).
FT MOD_RES 727 727 PHOSPHORYLATION (BY SIMILARITY).
FT SEQUENCE 749 AA: 87197 MW: 249919952B65F1 CRC64;

Query Match 49.2%; Score 1984; DB 1; Length 749;
Best Local Similarity 51.0%; Pred. No. 6,8e-117;
Matches 390; Conservative 141; Mismatches 205; Indels 28; Gaps 11;

OY 1 MAOMNLOQDTRYLKOLHLYSDTFPELROFLAPWIESQDAMAYAAKSKSHATLVFHNL 60
DB 1 MSQWELFQDLSKREGLVHLYSDTFPELROFLAPWIESQDAMAYAAKSKSHATLVFHNL 60
OY 61 LGELDOOYSRFLOESNVLYOHNLRIKQFIOSRYLEKPMETARIIVARCLMEESRLQTA 120
DB 61 LSQLDDYSRFSLENNFLLOHNIRKSRNLQDNFEDPVOMSMIYVCLKEERKILENAQ 120
OY 121 --TAAOOGGANHTAIVVTEKQOMLEQHLQDYARKRYQDLEQKKAYVENLODDDFVKT 178
DB 121 RPNDOAGNIGN----TWLMDKQELDSKVRNVAQDWCIEQELKTELEDEYDFCKT 176
OY 179 LKSGQDMQDINGNOSVYTRQKMODLEQMLTALDQMRISYSELAGLSAMEYVQKLTDE 238
DB 177 --SONRGEANGVAKSQDQEQLLHKMFMLDMKKREIHKIRELLNSTELQNTLIND 234
OY 239 ELADMKRRELACIGRPNIGLORLEWITSLSAESQLTQOUKLELQOKVSYKGDPI 298
DB 235 ELVEMKRRQOSACIGGPNACLDQLTQWFTIVAETLQIQOLKLELEQKFEYEDPI 294
OY 299 VOHRPMLEERIVELFRMLKSAFVYRQOPMPHMDRPLVIKIGVQTTVRLVLRPEL 358
DB 295 TKKKOVLSDPTFLPFOOLIOSSVYERQOPMPHMDRPLVKTGVQTVYVSRLLVLRQES 354
OY 359 NYOLKIKVCIKDKSGDVAALRGSRKFNILGTNTKVMNMBESNGSLSAEFRHHLTLEQRC 418
DB 355 NLLTKVCHPDKDVNENKTKGFRKFNILGTHTKVMNMBESNGSLSAELRHILQKQOK- 413
OY 419 GNGGRANCDASLIVTEHLITFETEVYHOGKIDLETLSLPVVVINSIQMPNMAASIL 478
DB 414 -NAGNRTNEGPLVTEHLISFETQCGPLVIDLETLSLPVVVINSVQSLSGMASIL 472
OY 479 WYMLLTNPKNVNFETPRPTGTDQVAVLSMOFSSSTTKRGLSIEOLJTTLAEKILGPGVN 538
DB 473 WYMLLTNPKNVNFETPRPTGTDQVAVLSMOFSSSTTKRGLSIEOLJTTLAEKILGPGVN 532
OY 539 YSCGQITWAKCEKKNAGKGSFWVWLDNIIDLVKYYIALNMEGYIMGFSIKERERAIL 598
DB 539 YSCGQITWAKCEKKNAGKGSFWVWLDNIIDLVKYYIALNMEGYIMGFSIKERERAIL 598

DB 533 PDGL-IPWTRCKENINDKNFSFWPMDITLIELIKNDLLCLMNGCIMGFISRKERALL 591
OY 599 STKPPGFTLLRFSESSREGVYTFWVENKDIS-GKTQIQSVETPYTKOOLNNKSAEIIIMGY 657
DB 592 KDQOPGFLLRFSESSREGAITFTWVERSQNGDEDFHAYEPYTKKELSAVTFPDIIIRNY 651
OY 658 KIMATNLVSLVLYLPDIPKEAFGK-YCRP-ESQEHPEADGSAAPYTKTFICVT- 714
DB 652 KVMAENENPEPLKLYLPVNDIKDAFGKYYSRPAEAPPELDDPKRTGYIKTELISVSE 711
OY 715 --PTTGSNTID-LPMSPTLDSLMQFGNNGCAEFSAGOFESL 755
DB 712 VHRPRLQTTDMLDLPMSPEPDEMARI-----VQPEPDSM 745

RESULT 6
STA4 HUMAN STANDARD; PRT; 748 AA.
ID ST4A HUMAN STANDARD; PRT; 748 AA.
AC 014765;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 4.
GN STAT4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Xu X., Sun Y.L., Hoey T.;
RT "The STAT amino-terminal domain mediates cooperative DNA binding
RT and confers selective sequence recognition."
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CARRIES OUT A DUAL FUNCTION: SIGNAL TRANSDUCTION AND
CC ACTIVATION OF TRANSCRIPTION.
CC -1- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
CC MEMBER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION.
CC -1- PTM: TYROSINE PHOSPHORYLATED. SERINE PHOSPHORYLATION IS ALSO
CC REQUIRED FOR MAXIMAL TRANSCRIPTIONAL ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L78440: AAB05605.1: -.
DR MIM: 600558: -.
DR InterPro: IPR000980: SH2.
DR InterPro: IPR001217: STAT.
DR Pfam: PF00017: SH2: 1.
DR Pfam: PF01017: STAT: 1.
DR SMART: SM00252: SH2: 1.
DR PROSITE: PS50001: SH2: 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW Phosphorylation; SH2 domain.
FT DOMAIN 569 664 SH2.
FT MOD_RES 663 693 PHOSPHORYLATION (BY JAKS) (BY
FT SIMILARITY).
FT MOD_RES 721 721 PHOSPHORYLATION (BY SIMILARITY).
FT SEQUENCE 748 AA: 85940 MW: 11E43803A9AF4FA CRC64;

Query Match 44.6%; Score 1798; DB 1; Length 748;
Best Local Similarity 48.4%; Pred. No. 3e-105;
Matches 357; Conservative 146; Mismatches 209; Indels 26; Gaps 10;
```

```

OY 1 MAOMNOLQDLTRFKYKOLHOLYSDTFPMELOF LAPMIESQDMAVAASKESHATLVFNHL 60
   1 MSQMNVOVQOLEIKFLEVOYDFDNEFMEIRHLAQMIENQDWEASNNETMATILLONL 60
OY 61 LGETIDQYGRFLOESNVLYOHNLRIKQLOFQSYLEKPMETIARIVARCLMEESRLQTA 120
   61 LIQLEIDQIGVSKENKLLIHNKLRIRKYLQGFHGNPMHVAVAVISNCLREBRIL-AAA 119
OY 121 TAAQGGQANHPPTAAVYTEKQMLFQHLQDVRKRYODLEQMKVENVLDDPFNFYKTLK 180
   120 NMPYQGLEKSLDSSVSEQSRVNEHKVAIKNSVQMTEDDTYLEDLODFEYKRYKTIQ 179
OY 181 SOGDMODLNGNNSVTRQKMOLEQMLTALDQNRISIVSELAQLASMEYVQKTLDEEL 240
   180 T----MDOSDKNSAMVNOEVLTLQEMLNLSLDFEKRKALSMTQIIEHTDLMTMLLEEL 235
OY 241 ADKRRPELACIGGPPNICDRLNENMTLSAESQLOTRQOIKKLELELOQVSYKGPITYQ 300
   236 QDKRRQOJACIGGPIHNGLDLOQNCFTLLAESLPOLRQLEKLEEDSTMTTEGDIPTM 295
OY 301 HRPMLEERIVELFRNLMSAFVVEROPCMHPDRPLVIKTVQFTTKVRLVKFPELNY 360
   296 QRTHMLERIVTEFLYNLFKNSFEVEROPCMPTHPQRPDLVKTLLQFTYKRLLIKIPELNY 355
OY 361 QLKIKVCIDKSDVAALQSKRFNLGNTKYNNNEESNGSLSAEFKLLTLEORCGN 420
   356 QYVVKASIDK---NVSTL-SNRPFVLCGNVVKAMSIEESSNGSLSEFRRHLOREKMSA 411
OY 421 GGRANDASLIYEEHLITFEFEVYHOGKIDLETHSLPVVVISNCOMPNMASTLWY 480
   412 GGGGN-EGCHMVTGEEHLSTTFETQICLYGLTIDLETSLPVIWISVQSLPNMASTIYW 470
OY 481 NMLTNPKNVNFFTKPPIGTMDQVAEVLWQFSSTTKRGLSIEQLTTLAEKLLGPVNS 540
   471 NVSTNDSQNLVFFNNPPPALISQLEWMSQFSYVGRGLNSDQLHMLAEKLTYS- SYS 529
OY 541 GCQITAKCKEMAKGSPFWMLDNIIDLYKYLALNEGTYGFIKEKEBALST 600
   530 DGHUTMAKPKCKEHLPKSPFTFWMLDLKIKKHLPLMIDYVGVGFSKEKERLLKD 589
OY 601 KPPGTLRFRESSESGGVTFWVEKDISGKTQIOSVEPTYTKOOLNMSAEIIMGYKIM 660
   590 KMGFTLLRSESHL-GGIFTWVHSESEGEVHFHSEVPYKNGRLSLAPADLLROKVI 648
OY 661 DATNLIYSLPLYTPDIPEKEAGKTCRPE---SOEHPADGSAAPYKTKFICVPTPT 716
   649 MAENLPEENLKYLPDIPKDAFKGHYSQPCVSRPTERGDKG---YVPVFIPISTI 704
OY 717 TCSNTLI-----LPMSP 728
   705 RSDSTEPHSPDLLPMSP 722

```

```

RT activators of transcription.*;
RN Proc. Natl. Acad. Sci. U.S.A. 91:4806-4810(1994).
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE-Testis;
RX MEDLINE-94277038; PubMed-8007943;
RA Yamamoto K., Quelle F.W., Thierfelder W.E., Kreider B.L.,
RA Gilbert D.J., Jenkins N.A., Copeland N.G., Silvennoinen O.,
RA Ihle J.N.;
RT Stat4, a novel gamma interferon activation site-binding protein
RT expressed in early myeloid differentiation.*;
RL Mol. Cell. Biol. 14:4342-4349(1994).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.45 ANGSTROMS) OF 1-123.
RX MEDLINE-98128033; PubMed-9461439;
RA Vinkemeier U., Moareff I., Darnell J.E. Jr., Kurian J.;
RT Structure of the amino-terminal protein interaction domain of
RT STAT-4.*;
RL Science 279:1048-1052(1998).
CC -1- FUNCTION: CARRIES OUT A DUAL FUNCTION: SIGNAL TRANSDUCTION AND
CC ACTIVATION OF TRANSCRIPTION.
CC -1- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
CC MEMBER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION.
CC -1- PTM: THROSYNE PHOSPHORYLATED. SERINE PHOSPHORYLATION IS ALSO
CC REQUIRED FOR MAXIMAL TRANSCRIPTIONAL ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 006923; AAA19453.1; -.
DR DDB: 009351; AAA19692.1; -.
DR PDB: 1BGF; -SEP-98.
DR TRANSFAC: J01576; -.
DR MGD: MGI:104062; Stat4.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00...42; 1.
DR SMART: SM...TAT; 1.
DR ProSite: P30001; SH2; 1.
DR Transcription regulation: DNA-binding; Nuclear protein;
KW Phosphorylation; SH2 domain; 3D-structure.
FT DOMAIN 570 665
FT MOD_RES 694 694
FT PHOSPHORYLATION (4*) (K)S (BY
FT SIMILARITY).
FT MOD_RES 722 722
FT PHOSPHORYLATION (4*) (S) (BY
FT SIMILARITY).
FT CONFLICT 198 638 MISSING (IN REF. 2).
FT CONFLICT 198 638 A -> P (IN REF. 2).
SQ SEQUENCE 143 AA; 85940 MW; A88B837E49CFEBC CRC64;

Query Match 44.2%; Score 1780.5; DB 1; Length 749;
Best Local Similarity 48.1%; Pred. No. 3.8e-104;
Matches 356; Conservative 149; Mismatches 206; Indels 29; Gaps 11;

```

```

Db 120 NMPDQPLEKSLQSSSVSEQRNVEKVAIKKSVQMTEDDTYLYEDLOEDFYRTKTIQ 179
Oy 181 S--OCDMODLNGNNOSTROKMOOLEOMLTALDOMRSIYSELGLASMEYQKTLTDE 238
Db 180 TMDGOK-----NSILVNEVLTLLDOLMNSLDPKRKREALKSKTOIVNEFDLMSMLE 234
Oy 239 ELADWKRREPLACIGPPNICLDRENMWITSAESLOLRQOIKLEELQOKVYKGDPI 298
Db 235 ELQDMKRRQOICIGPRLHNGDLOQNCFTLLAESL.FOLROLEKLEQOSTKMTYEGDPI 294
Oy 299 VOHRPMLERIEVELFRNLKKSATVVERQRCMPHPRPLVIKRGVOFTTKVLLKPEL 358
Db 295 PPARHLLEPAFLYLNLFKNSFVERQPCMPHPRPMLVKTLLQFTVYRLRLIKLPEL 354
Oy 359 NVOLKIKVCIKDSQVVALRGSRKNIIGTNTKVNMMESNNGSLSAEKKHLTLRQRC 418
Db 355 NVOVKKASIDK---NVSTL-SNRREVLCGTHVKAASSSESSNGSLSEVERHLQPKEMKC 410
Oy 419 GNGGRANCASLIYTEELHLTFETEVYHOGKIDLETHSLPVVVISNICQMPNANASIL 478
Db 411 STGSKCN-EGCHVYTELHSTTFETOICLYGLTINLETSSLPVVVISNVSQLPNANASII 469
Oy 479 WYNMLTNKNNKNNFEKPPRGITMDQVAEVLMSQSSSTTKRGLSEDTLTAELKLGPNV 538
Db 470 WYNVSTNDSONLVFNNPSSVTLGQLEVMASWQSSYVGGLSEQDLMNAELKLVQOS-N 528
Oy 539 YSCQITMAKFKCKENNAAGCFSEFWALNDIIDLVKRYILALMEGYIMGPISEKEREAIL 598
Db 529 YNDGHJLTMKFKCEHLPGKTFETFWLLEAILDLIKKHILPLMDGYIMGVSKKEKELL 588
Oy 599 STKRPCTFLIRSESEKSGCVFTWYKEDIQIOSEVPYTKOOLNNMSFELIIMGYK 658
Db 589 KKMPCFTFLIRSESHL-IGITFTWDOSENGEVREHSEVPYKGRISALAFADIIIRDYK 647
Oy 659 IMDATNITLSPYLYLPDIKPEAFGKYCRPE-----SOEHPEADPGSAAYLTAKFCVPT 714
Db 648 VYMAENIPENPLKYLTPDLPKAKAFKHHSSQCEVSRPTERDCKG-----YVPSVFILPIS 703
Oy 715 PTTCSMTID-----LPMSPT 728
Db 704 TIRSDTEPQSPDLPMSP 723

RESULT 8
STA2_HUMAN STANDARD: PRT; 851 AA.
AC P52630: Q16430: Q16431:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 2 (P113).
GN STAT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95192056; PubMed=7885841;
RA Yan R., Qureshi S., Zhong Z., Wen Z., Darnell J.E. Jr.;
RT "The genomic structure of the STAT genes: multiple exons in
RT colicident sites in Stat1 and Stat2.";
RL Nucleic Acids Res. 23:459-463(1995).
RN [2]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE=93366558; PubMed=1502204;
RA Fu X.-Y., Schindler C., Impirota T., Abersold R., Darnell J.E. Jr.;
RT "The proteins of ISGF-3, the interferon alpha-induced transcriptional
RT activator, define a gene family involved in signal transduction.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:7840-7843(1992).
RN [3]
RP ALTERNATIVE SPLICING.

```

```

RX MEDLINE=96176320; PubMed=8601453;
RA Sugiyama T., Nishio Y., Kishimoto T., Akita S.;
RT "Identification of alternative splicing form of Stat2.";
RL FEBS Lett. 381:191-194(1996).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=97172544; PubMed=9020188;
RA Bluyssen H.A., Levy D.E.;
RT "Stat2 is a transcriptional activator that requires sequence-specific
RT contacts provided by stat1 and p48 for stable interaction with DNA.";
RL J. Biol. Chem. 272:4600-4605(1997)
CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE IFN-STIMULATED
CC RESPONSE ELEMENT (ISRE) AND TO THE GAS ELEMENT. THIS MULTIPROTEIN
CC TRANSCRIPTION FACTOR IS TERMED ISGF3.
CC -1- SUBUNIT: IN RESPONSE TO IFN ALPHA/BETA, THREE SUBUNITS (STAT1-
CC ALPHA, STAT1-BETA, STAT2) OF ISGF3, BECOME PHOSPHORYLATED ON
CC TOGETHER WITH ISGF3 GAMMA (P48), A DNA-BINDING PROTEIN THAT
CC SPECIFICALLY BINDS TO THE IFN-STIMULATED RESPONSE ELEMENT.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC -1- RESPONSE TO PHOSPHORYLATION.
CC -1- SHORT FORM: ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IFN-ALPHA.
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U18671; AAA98760.1; -
DR EMBL: M97934; -; NOT_ANNOTATED_CDS.
DR EMBL: S81491; AAB36226.1; -
DR EMBL: S81491; AAB36227.1; ALT_SEQ.
DR HSSP: P42224; 1BF5.
DR MIM: 600556; -
DR InterPro: IPR000980; SH2
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW Phosphorylation; SH2 domain; Alternative splicing.
FT DOMAIN 572 667
FT MOD_RES 690 690
FT VARSPIC 621 652
FT FT
FT FT
FT VARSPIC 653 851
FT FT
FT SEQUENCE 851 AA; 97916 MW; E4C74674CB7A3215 CRC64;

```

```

Query Match 32.1%; Score 1293; DB 1: Length 851;
Best local similarity 38.6%; Pred. No. 166-73;
Matches 298; Conservative 148; Mismatches 276; Indels 50; Gaps 18;

Oy 1 MAQWNLQQLDTRYLKQLHQLYSDT-FPEMLRQFLAPWIESQDMAYAA--SKESHATLVF 57
Db 1 MAQWENLQNLDSFQDLHQLYSHSLPVDIRYLAVWIEDQWQAAAGSDSKATIMF 60
Oy 58 HNLGIIIDQYSEFLDE-SVLYQVHMLRKQFLQSGRYLSEKPEIARIYARCLMEESRL 116
Db 61 FHLDLQNLWECGRCSDPESLLDQHLNLRKFCRDQIP-FSQDPTQLAEMIFNLLEKRL 119
Oy 117 QYATATV AVOGGGQANHPPTAAVVTPEKQO-MLDQHLQDVRKRVQDLEQMKVVENLDQDFDN 175

```


Db	284	SLFRVQGMFAAGVDDLBRNQWELLQRLRLDRSEFVETQCPMQOTLHRLPILLTGKKFTFR	343
Oy	349	VALLKFPFLNATOLKIKVCIDKQSCDVAALRQSKRFNLIIGTITKYMNNEESNNGSLAEF	408
Db	344	TRLVLRLQEGSSSLAEVSDRNSD---LPGFRKFNLTISNOQKLTLPKCGORGLIMDF	399
Oy	409	KHLTREORGCGNRCANCSALIVTEELHLIFFEEFVNYHQGLKIDLETSLPVVYISNIC	468
Db	400	GLTLVEQRAVAGAGKGNKKNGPLATVTEELHIVSFVETTYOGLKMKLOTDITLPPVYIISNN	459
Oy	469	QMPNMASTILWYMLTNPNKRVNFFTKPEIGTWDAEVLASQFSSTTKRGLSDLTLL	528
Db	460	QLSPAMASTILWMLSPNKNQOFCQAPKAWSLGLPVLISQFSYSYARGLDSEGLML	519
Oy	529	AEKLLGPGVNSGCOITNAKPKCK-ENMGCKGFSFWYMLNDIIDLKYYILALMNGEYING	587
Db	520	KTKLKGSKCKMEDALLSWDFCKRESPECK-IPFWMLDKLLELVDHLKLDLCKDRING	578
Oy	588	FISKEREALISTKPCPTFLRFSSESSKEGCVTFTWVKEDISGKTQIOSEVPEYTKQOLNN	647
Db	579	FVSRQDERLLKKMLSGFTLRFSEFS-EGGITCSWVEHQDDHKVEIYSVGQYTKLVLDS	637
Oy	648	MSFAELIMKTKIMDATNIIIVSLVLYLPIDIPKEEAFGKCRP-----ESEHNEADPGSAA	703
Db	638	LPLETLIRYVOALAEENIPENPLRFLYLPRIPEDEAFGCVYQKRVMLDEQE-----	688
Oy	704	PLYLTKFCIVPTPTCSNTIDLPMSPTDLQSLMQ	736
Db	689	YLKHKHLIV-----SNRQVDELQ	707
RESULT	10		
STA2_PIG	ID	STA2_PIG	STANDARD; PRT; 864 AA.
AC	002799;		
DT	15-JUL-1999 (Rel. 38, Created)		
DT	15-JUL-1999 (Rel. 38, Last sequence update)		
DT	15-JUL-1999 (Rel. 38, Last annotation update)		
DE	SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 2.		
GN	STAT2.		
OS	Sus scrofa (Pig).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.		
OX	NCBI_TaxId=9833;		
KN	(1)		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Muscle;		
RA	Ito Y., Mikawa S., Kobayashi E., Wada Y., Minezawa M.;		
RL	Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.		
CC	- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE IFN-STIMULATED		
CC	RESPONSE ELEMENT (ISRE) AND TO THE GAS ELEMENT. THIS MULTIPROTEIN		
CC	TRANSCRIPTION FACTOR IS TERMED ISGF3 (BY SIMILARITY).		
CC	- SUBUNIT: IN RESPONSE TO IFN ALPHA/BETA, THREE SUBUNITS (STAT1-		
CC	ALPHA, STAT1-BETA, STAT2) OF ISGF3, BECOME PHOSPHORYLATED ON		
CC	TYROSINE, MIGRATE INTO THE NUCLEUS, AND ASSEMBLE INTO A COMPLEX		
CC	TOGETHER WITH ISGF3 GAMMA (P48), A DNA-BINDING PROTEIN THAT		
CC	SPECIFICALLY BINDS TO THE IFN-STIMULATED RESPONSE ELEMENT (BY		
CC	SIMILARITY).		
CC	- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN		
CC	RESPONSE TO PHOSPHORYLATION (BY SIMILARITY).		
CC	- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IFN-ALPHA (BY		
CC	SIMILARITY).		
CC	- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.		
CC	- SIMILARITY: CONTAINS 1 SH2 DOMAIN.		
CC	-----		
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		

CC	EMBL: AB004061: BAA20332.1: -	Score 1264.5; DB 1: length 864;
DR	InterPro: IPR000980: SH2.	
DR	InterPro: IPR01217: STAT.	
DR	Pfam: PF00017: SH2; 1.	
DR	Pfam: PF01017: STAT; 1.	
DR	SMART: SM00252: SH2; 1.	
DR	PROSITE: PS00001: SH2; 1.	
KM	Transcription regulation; DNA-binding; Nuclear protein;	
KM	Phosphorylation; SH2 domain.	
FT	DOMAIN 572 667 SH2.	
FT	MOD_RES 590 690 PHOSPHORYLATION (BY JAKS) (BY SIMILARITY).	
SO	SEQUENCE 864 AA: 99060 MM: 28FCG9569F5B20 CRC64:	
Query Match	31.4%; Score 1264.5; DB 1: length 864;	
Best Local Similarity	37.6%; Pred. No. 9,9e-72;	
Matches 292; Conservative 149; Mismatches 279; Indels 57; Gaps 18;		
QY	1 MAMWMDLOOLDTRYLKOLHOLYSDF--PPELTKROPFLAPWTEESODMAVAA--SKESHATVLF 57	
DB	1 MAMWMDLNDSPFODOLQOLTSLSLPADVQYLAWTEEDQWDBAALGNDSCRANMLF 60	
QY	58 HNLLGEIDQYSRFLDESN-VLYQHNLRKIKOFLOSRYLEKPMELARIVAKRLTMEESRL 116	
DB	61 FHLRLQNLVDCGRGDRECLLDNLRKFRYDIA-IPOGRPLRAEMI-FNLLLEKRL 119	
QY	117 QTAATAAOOGGANIPTAAVTEKQ-MLEQNLQDYRKVVQDLQKKKVVENLQDDPDN 175	
DB	120 IQAQRAQL--EQDRALEARGENQHNLESLLELRAMEKLVKLSISLQKQDQIFCR 176	
QY	176 YKTLKSGDMQDLNNGNSVTRKKMOOLEOMLTALDMKRSIVSELAGLSAMEVYUKTL 235	
DB	177 YK-IKASAKTHSLDHR---TRQE-QVLOETLNELDKRRKEVLDASKALGLKTLTLEL 231	
QY	236 TDEELADMKRPELACIGGRPNICLDRLNNVITSLAESQLQTRQOIKLEELDQKVSXKG 295	
DB	232 L-PKLEEMKVVQOKAKACIGAPMDGELEQLEKMTAEAKLFLNLRQLLEIKLGSVVKYDE 290	
QY	296 DPIVQHRPLLEIRVIELFRNLKMSAFVVERQSCMRPHRRIVITGVQVTTKVRILAKF 355	
DB	291 DLLEFGVDLTKQVTELLQRLRHRAFIYETQRCMOTQPRRLRLTKGSGFVTRLLVNL 350	
QY	356 PELANYOLKIKVCIDKDSGDVAALKGSRKFNILGTNTKYVNMEESSNGSLSAEFKHILTRE 415	
DB	351 QGNSGTIAFVSIIRNP---PKSGFRKFRNLITSNKTLTPKGGSGQGLWPGVLTLE 407	
QY	416 QRCGNGRANCDA SLVFEELHLITFEFVYHQGLKIDLETSLVYVVISNIGCPNAMA 475	
DB	408 QRSGGAGGNNKGRPLGVTEELHISFTVKKYTOGKQELTPTTLPLVVISNNQSLIAMA 467	
QY	476 SLIWNMLTNNKNNVNFETKPRIGTWQDVAEVLISWQFSSTTKRGSLIEQLTDLAEKLGPR 535	
DB	468 SLIWNMLTNSDPQNOQGFSSPRKAWNLLGRALSMQFSHWQGLNSDQGLMRKRLQ 527	
QY	536 GVNISGCOITMAKFK-ENMAGKGSFWWVLDNIIDLKYLALMNEGVIINGTFSKERE 594	
DB	528 NSSTGLSLSWDFIKRSPQCK-LPFWTMTLHLDLVHDLHKLDMKQGHINGVSRSE 586	
QY	595 RALSTKPRGFFLAFSSSEKGGVTFWVMEKDIGKTIQISVEVRYTQQLNNMSFAELI 654	
DB	587 RRLKTKTISGTFILRFSE-TLEGGCTGCVWEHQDDDKYALISLQPTKEVLOSLPLTII 645	
QY	655 MGKYMADATNLLVSLVLYVYDPIDPEEAFGKKYCRPESQENHPADSGAARYLKTGFICVT 714	
DB	646 SOYQLLTLEENIPENDRLRFLYPRIPDEAFGC---NDE-KANQDEKKYKILKHLIYVS 699	
QY	715 -----PTTCSNTIDLPMSPRLDSLMQFNGNGBGAEPSAGQFESL 755	
DB	700 NROYDELDQPEPLKLEPLLESLDLDGLAP-----GPBQVGLDLEPL 742	


```

RL Proc. Natl. Acad. Sci. U.S.A. 92:8831-8835(1995).
CC -i- FUNCTION: CARRIES OUT A DUAL FUNCTION: SIGNAL TRANSDUCTION AND
CC ACTIVATION OF TRANSCRIPTION. BINDS TO THE SIGNAL ELEMENT AND
CC ACTIVATES PRL-INDUCED TRANSCRIPTION.
CC -i- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
CC MEMBER (BY SIMILARITY).
CC -i- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION.
CC -i- TISSUE SPECIFICITY: IN THE VIRGIN, FOUND IN MOST TISSUES.
CC PARTICULARLY ABUNDANT IN MUSCLE TISSUE OF VIRGIN AND LACTATING
CC FEMALES, AND OF MALES.
CC -i- DEVELOPMENTAL STAGE: DETECTED BOTH IN VIRGIN MOUSE AND AFTER
CC MAMMARY GLAND INVOLUTION. THE LEVEL OF STAT5A INCREASES CONSTANTLY
CC DURING PREGNANCY, BUT DECREASES DURING LACTATION.
CC -i- PTM: TYROSINE PHOSPHORYLATED.
CC -i- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -i- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: Z48539; CA88420.1; -
DR EMBL: U21110; AAC52282.1; -
DR HSSP: P42224; 1BP5.
DR TRANSFAC: T00944; -
DR MGD: MGI:103035; Stat5b.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS0001; SH2; 1.
DR Transcription regulation: DNA-binding: Nuclear protein:
DR KW phosphorylation; SH2 domain.
DR FT DOMAIN 589 686 SH2.
FT MOD_RES 699 699 PHOSPHORYLATION (BY JAKS) (BY
FT SIMILARITY).
FT CONFLICT 433 433 E -> G (IN REF. 2).
FT SEQUENCE 786 AA; 90002 MW; A8FE76405E41B2EF CRC64;

Query Match 21.5%; Score 865.5; DB 1; Length 786;
Best Local Similarity 30.1%; Pred. No. 8,6e-47;
Matches 247; Conservative 146; Mismatches 317; Indels 111; Gaps 30;

OY 1 MAOWNLOOLDTRYKOLHQLQSDPFPMELROFLAPWISSQOMAYV-----ASKESHATLV 56
DB 1 MAMWIOAOOLQGDALHOMQALIGOEPIEVRYHLSOWISQAMDSIDLDPQENINATOL 60
OY 57 FNNHLSGEIDQOYSRFLQESNVLYOHNLRIKQFLQSRYLEKPMETARIYARCLMESRLL 116
DB 61 LSGIVGELQKRAKHQVGEGLFKIKLGHYATOLQSTVYRCMELVRCRLHILYNQORLY 120
OY 117 QTAATRAOOGGAGNHPTAAVTEKQOMLEQHLQDVYKRVQDLQKKKVVENLQDDFFNY 176
DB 121 REANNSSSPAGS-----LADAMSQKHQIQTQFEELRLITQDENELKRLQOQOEYFIQY 176
OY 177 -RTLSQSGQMDLNGNN-----OSVTRQKKMOOLEQML-----TALDQRRSTVSELAL 224
DB 177 QESLRIOAQFOALQOLNPOERSRSTALQOKOVSLQETWLOREAQTLQOYRVELAEKHQRT 236
OY 225 LSAMEYVQKTLTDEELADKKRRPELACIGPPNICIDRLNNITSLAESQLOTRQOIKL 284
DB 237 LQLLRKQQTIIILDELQIKKRRQQLANGNGPREGSLDVLOSCEKLAELIIMONROQIRRA 296
OY 265 BELQOQVSKGPIVOHQRPMLEERIVLEFRNLKMSAFVVEROPCPMDRPLVIRKTVQ 344
DB 297 EHLCOOLPIFG-PVEEMLAEVNATITDLSALVTSTFIIEKOP-----POVLKTOYRK 347

```

```

OY 345 FTTKRYELVFPPELNYOL---KIKVCIDKDSGVAAIKRSRKN-----ILGTRKYNNM 396
DB 348 FAATVRLVG-GLNVHMPPOVKATITISQQAQKSLKKNENTRNDYSGEIL--NMCVMEY 405
OY 397 EESNNGSLSAEFKHLTLRQRCNGCRANDASLIYTEELHLITFETEVYHQC--LKIDL 454
DB 406 HOA-TQTLNHFERNMSLK--RIKRSRGAES--VTEKFTTLTFDSQFVSGQNELVQV 459
OY 455 EFTSLPWWYISNIGQPNMASILWYNNMLTNPKKNVNFPTPIGWDAVAVLSWQFS 514
DB 460 KTLSPVWYIVHSQNNATATVLMNARA-EGKVP-FAVPRKVLMPQLCEALNNKRYKA 517
OY 515 --TTKRLSIEDLTTLAEKILGPNV---YSCQITMAKFKENNAGKGFSTWVLQNI 568
DB 518 EVQSNNGLTKEINLVFLAQKLFNISNHLLEDYNSVSWQSFQNFRENIPLGRNYTFWQMDGV 577
OY 569 IDLVKRYIILALWNEGYIMGYSKEREKRALISTPPTFTLLRSESSKGGVFTWYEXI 628
DB 578 MEVLKHLKPRHMDGAILGFWNKQAHDLLINKPDQTFILRSD-SEIGQITLAW--KFD 634
OY 629 SGKTOISVEPYTKQOLNNKSPAEIIMGYKIMATNIIYSPLYLYPDIPKEAFQKY-- 686
DB 635 SQERMWNLMPFTTRDFSIKSLAD-----RLGD-----LNYILYVPRPKDEVYSKYVT 684
OY 687 ---CRPES-----QHPPE-----ADPGSAAPYLKTKFLCVIPPTQSNITDLP 725
DB 685 PVCEPATAKKADGYVVKPOIKQVPEFANASTDAGSGATYMDQ---APSPVCPQA-IYN 740
OY 726 MSPRTLDSLMOFGNCGEAPSAQGFPSLTFPMDLTSECA 766
DB 741 MYPPNDSVLD--TDGD-----FDLEDITMDVA 765

RESULT 13
STSB_RAT
ID STSB_RAT STANDARD: PRT: 786 AA.
AC P52632;
DT 01-OCT-1996 (rel. 34, Created)
DT 01-OCT-1996 (rel. 34, Last sequence update)
DT 20-AUG-2001 (rel. 40, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 5R.
GN STAT5B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1] NCBI_Taxid=10116;
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=96102059; PubMed=8530402;
RA Ripberger J.A., Fritz S., Richter K., Hocke G.M., Lottspeich F.,
RA Fey G.H.;
RT *transcription factors Stat3 and Stat5b are present in rat liver
RT nuclei late in an acute phase response and bind interleukin-6
RT response elements *;
RL J. Biol. Chem. 270:29998-30006(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NOBLE; TISSUE=Lymph node;
RA Luo G., Yu-lee L.;
RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
CC -i- FUNCTION: CARRIES OUT A DUAL FUNCTION: SIGNAL TRANSDUCTION AND
CC ACTIVATION OF TRANSCRIPTION. BINDS TO THE SIGNAL ELEMENT AND
CC ACTIVATES PRL-INDUCED TRANSCRIPTION (BY SIMILARITY).
CC -i- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
CC MEMBER (BY SIMILARITY).
CC -i- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION (BY SIMILARITY).
CC -i- PTM: TYROSINE PHOSPHORYLATED (BY SIMILARITY).
CC -i- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -i- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----

```


EMBL: U47686; AAC50491.1; -
 DR HSSP: P42224; 1BF5.
 DR MIM: 604260; -
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF00017; STAT; 1.
 DR Pfam: PF01017; STAT; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: P550001; SH2; 1.
 DR Transcription regulation: DNA-binding; Nuclear protein;
 KW Phosphorylation; SH2 domain.
 FT DOMAIN 589 686 SH2.
 FT MOD_RES 699 699 PHOSPHORYLATION (BY JAKS) (BY
 FT SIMILARITY).
 FT DOMAIN 232 321 REQUIRED FOR INTERACTION WITH NMI.
 FT CONFLICT 230 230 A -> P (IN REF. 2).
 FT CONFLICT 628 628 S -> T (IN REF. 2).
 FT CONFLICT 717 717 D -> DA (IN REF. 2).
 FT CONFLICT 720 720 R -> G (IN REF. 2).
 SQ SEQUENCE 786 AA: 89880 MW: 84265C4C8EB824B6 CRC64:

Query Match 21.4%; Score 862.5; DB 1; Length 786;
 Best Local Similarity 30.2%; Pred. No. 1.3e-46;

Matches 248: Conservative 151; Mismatches 311; Indels 111; Gaps 31;

OY 1 MAQNNQLOQDTRYLKQHLQYSDTFPMELRQFLAPWIESQDWAYA---ASKESHATLV 56
 DB 1 MAVVIAQAOLOGEALHOMALYQHFIEVRHYLSQWIESQAMDSYDLDPENIKATQL 60
 OY 57 FNNLIGETDOQYRFLQESNVLYQHNLRIRIKQFLQSRYLEKEMELARIARCLMESRL 116
 DB 61 LEGVQELQKKAEHOGVGEQGLTKLGHATQLONTYDRCMEVRCRHLILYNNQRLV 120
 OY 117 QTAATAQGGQGANHTAAYVTEKQMLEOHLQDYRKRQVLEQKKVVENLQDDDFNY 176
 DB 121 REANNSSPAGS---LADAMSQKHLQINQTFEELRLVYQDENELKRLQOQEFIIQY 176
 OY 177 -KTLKQ--GDMQNLG---NNQSVTRQKKQOLEQML---TALDQRRSIVSELAQL 224
 DB 177 QESLRIQAFQGLQSLQFQERLSQERLQALQKQVSLQEAQLEQRLQOQRYVLEAKHQKT 236
 OY 225 LSAMEYVQKTLDELQADMKRREPLACIGGPPNICLDRLQENITSLAESQLOTRQIKKL 284
 DB 237 LQLLRQKQTIIDDELQMKRQKQALAGNGPREGSLVQSCERLALITIMONRQOIRRA 296
 OY 285 BELQOKVSGDPIVQHRPMLLEERIVLEFRLMKSAFVEROPCMHPRDRPLVITKGYO 344
 DB 297 EHLQCOIHPFG-PVEEMLAENVATITDIISALVSTFIIERQK-----POVLKQTRK 347
 OY 345 FTTKRLVLPKPELNYQ---KIKYCIKQSDQVAALGSKRFN-----TLCTNTRKVMNM 396
 DB 348 FAATYRLLVG-GKLNVHNNPQVQKATIIISQAKSLKNENTRNDYSGEIL--NNCCVMEY 405
 OY 397 EESNNGSLAEFKHLTLREORCGNGRANCASLIIVTEELHLITPEVEYHOG--LKLIDL 454
 DB 406 HQA-IGTLSAHRNNSLK--RIKRSDRGAGES---VTEBEKTLIFESQPSVGCNELVROY 459
 OY 455 ETHSLPVPVVISNICMPNMAASILMYNMLTNNPKVNFETKPPIGTWDQVAEVLISWOF 514
 DB 460 KTLSPVAVVIVHGSQDNNATATVLDNAFA-EPGRVP-FAVPDKVLMPQLCEALMKRKA 517
 OY 515 -TTRRGISTEQLTLAKKLIGPG---VNSGCOITWAKCKCKEMAKGGSFVWVLNDI 568
 DB 518 EVQSGRGLTKENLVFLAQLKLNSSSHLEDYSGLSVNSQSNRENRLPGRNTYTFWQMPFGV 577
 OY 569 IDLVKKYLLALMNGYINGFTSKEREKRALSTKPPGTFLLRFSSSKGCVFTVWEKDI 628
 DB 578 MEVLAKKHLKPHMNGALIGFVNKQAHDLINKPDGTFLLRFSQ-SELGGSIAM--KFD 634
 OY 629 SGKQIQQVSEYRTKQQLNNMSFAELIMGYKIMDATNIVSPVLYLPDIPKEEAFGKYCR 688
 DB 635 SQERFMFNMMPFTTRDSIRSLAD-----RLGD-----LNVLIYVFPDRKDEVYSKYTT 684

OY 689 P---ES-----QEHPE-----ADPGSAAPYLTKPICVPTTCSTNTIDLP 725
 DB 685 PVPCEATAKKAVDGYVKKPQIKQVYPEFVNASADGCGSATYTMQ---APSPAVCPQA-HYN 740
 OY 726 MSFRLIDSLMFGNNGCEGAPSAAGQFESLTFDMLTSCA 766
 DB 741 MYPQNDPSVLD--TDGD-----FDLEDTMDVA 765
 RESULT 15
 ST5A_HUMAN
 ID ST5A_HUMAN STANDARD: PRT: 794 AA.
 AC P42229;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 5A.
 GN STAT5A OR STAT5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=95236292; PubMed=7719937;
 RA Hou J., Schindler U., Henzel W.J., Wong S.C., McKnight S.L.;
 RT *Identification and purification of human Stat proteins activated in
 response to interleukin-2.*;
 RL Immunity 2:321-329(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Lin J.X., Nietz J., Modi W.S., John S., Leonard W.J.;
 RL Submitted, (EC-19c5) to the EMBL/GenBank/DBJ databases.
 CC -I- FUNCTION: CARRIES OUT A DUAL FUNCTION: SIGNAL TRANSDUCTION AND
 ACTIVATION OF TRANSCRIPTION. BINDS TO THE GAS ELEMENT AND
 ACTIVATES PRL-INDUCED TRANSCRIPTION.
 CC -I- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
 MEMBER (BY SIMILARITY).
 CC -I- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
 RESPONSE TO PHOSPHORYLATION.
 CC -I- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IL-2, IL-3, IL-7, IL-
 15, GM-CSF, GROWTH HORMONE, PROLACTIN, ERYTHROPOIETIN AND
 THROMBOPOIETIN. TYROSINE PHOSPHORYLATION IS REQUIRED FOR DNA-
 BINDING ACTIVITY AND DIMERIZATION. SERINE PHOSPHORYLATION IS ALSO
 REQUIRED FOR MAXIMAL TRANSCRIPTIONAL ACTIVITY (BY SIMILARITY).
 CC -I- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
 CC -I- SIMILARITY: CONAINS 1 SH2 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: L41142; AAA73962.1; -
 DR EMBL: U43185; AAB0589.1; -
 DR HSSP: P42224; 1BF5.
 DR MIM: 601511; -
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF01017; STAT; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: P550001; SH2; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein;
 KW Phosphorylation; SH2 domain.
 FT DOMAIN 589 686 SH2.
 FT MOD_RES 699 694 PHOSPHORYLATION (BY JAKS).
 FT CONFLICT 784 780 PHOSPHORYLATION (BY SIMILARITY).
 FT CONFLICT 88 88 G -> R (IN REF. 2).

